

VŠB – Technická univerzita Ostrava  
Fakulta elektrotechniky a informatiky  
Katedra informatiky

# **Modelování směrových dat pomocí bioinspirovaných výpočtů**

## **Modelling Directional Data by Bio-inspired Methods**

# Zadání diplomové práce

Student:

**Bc. Michal Šamaj**

Studijní program:

N2647 Informační a komunikační technologie

Studijní obor:

2612T025 Informatika a výpočetní technika

Téma:

**Modelování směrových dat pomocí bioinspirovaných výpočtů**  
**Modelling Directional Data by Bio-inspired Methods**

Jazyk vypracování:

čeština

Zásady pro vypracování:

Práce se zaměří na návrh nových bioinspirovaných algoritmů pro modelování směrových dat. Směrová data (např. směr větru) jsou charakteristická tím, že se jednotlivé body nacházejí na jednotkové kružnici a proto je třeba je modelovat jinak, než data lineární. Vedle optimalizace tradičních statistických modelů budou pro účel modelování směrových dat zváženy také metody strojového učení.

Práce bude řešena v následujících krocích:

1. Seznámení se s problematikou.
2. Vyhledání vhodných směrových dat.
3. Návrh algoritmu, implementace řešení.
4. Experimentální ověření, porovnání se stávajícími modely.

Seznam doporučené odborné literatury:

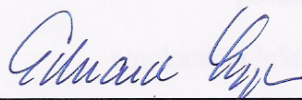
- [1] Masseran, N., Razali, A., Ibrahim, K., Latif, M., 2013. Fitting a mixture of von Mises distributions in order to model data on wind direction in Peninsular Malaysia. Energy Conversion and Management 72 (0), 94 – 102, the III. International Conference on Nuclear and Renewable Energy Resources {NURER2012}.
- [2] Heckenbergerova, J., Musílek, P., Mejznar, J., Vancura, M., 2013. Estimation of wind direction distribution with genetic algorithms. In: CCECE. IEEE, pp. 1–4.
- [3] Carta, J. A., Bueno, C., Ramírez, P., 2008. Statistical modelling of directional wind speeds using mixtures of von mises distributions: Case study. Energy Conversion and Management 49 (5), 897 – 907.

Formální náležitosti a rozsah diplomové práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

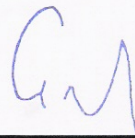
Vedoucí diplomové práce: **doc. Ing. Pavel Krömer, Ph.D.**

Datum zadání: 01.09.2015

Datum odevzdání: 29.04.2016



doc. Dr. Ing. Eduard Sojka  
vedoucí katedry



prof. RNDr. Václav Snášel, CSc.  
děkan fakulty



Prehlasujem, že som túto diplomovú prácu vypracoval samostatne. Uviedol som všetky literárne  
pramene a publikácie, z ktorých som čerpal.

V Ostrave 20. apríla 2016

.....  
Samař

Rád by som na tomto mieste poďakoval pánovi doc. Ing. Pavelovi Krömerovi, Ph.D., ktorý ma vďaka plnohodnotným konzultáciám viedol k úspešnému zvládnutiu tejto práce. Rovnako rád by som poďakoval aj mojej priateľke Janke Odnogovej, ktorá ma veľmi podporovala a dodávala silu počas písania tejto práce.

## Abstrakt

Smerové dáta sú obsiahnuté v mnohých výskumných oblastiach. Rýchlosť a smer vetra sú najdôležitejšie premenné pre efektívne využitie veternej energie. Vietor tiež významne ovplyvňuje aktuálne prenosové kapacity elektrického vedenia. Preto je dôležité poznať ročné rozdelenie smeru vetra pre špecifické lokality, napr. v miestach kde sa nachádzajú veterné farmy alebo elektrické vedenie. V tejto práci je prezentovaná nová metóda pre určenie rozdelenia smeru vetra. Štatistický model je určený zmesou kruhovej von Misesovej distribúcie. Parametre modelu sa odhadujú pomocou vyhľadávacieho mechanizmu evolučných stratégií. Kvalita vypočítanej distribúcie sa vyhodnocuje pomocou Pearsonovho chi-kvadrát testu dobrej zhody alebo pomocou RMSE. Celá procedúra je potom testovaná pomocou prípadovej štúdie.

**Kľúčové slová:** von Misesova distribúcia, kruhová normálna distribúcia, evolučná stratégia, smer vetra, štatistické modelovanie

## Abstract

Directional data are included in many research fields. Wind speed and direction are the most important variables for the efficient use of wind energy. Wind also significantly influences actual carrying capacity of power transmission lines. Therefore, it is important to know the annual wind direction distribution for specific location, for example where wind farms or power transmission lines are located. In this thesis, a new method of wind direction distribution determination is presented. The statistical model is designed using circular von Mises distribution. Model parameters are estimated using search mechanism of evolutionary strategies. The quality of computed distribution is evaluated by Pearson's chi-squared test or by RMSE. The whole procedure is then tested using case study.

**Key Words:** von Mises distribution, circular normal distribution, evolution strategy, wind direction, statistical modelling

# Obsah

<b>Zoznam použitých skratiek a symbolov</b>	<b>8</b>
<b>Zoznam obrázkov</b>	<b>9</b>
<b>Zoznam tabuliek</b>	<b>10</b>
<b>1 Úvod</b>	<b>11</b>
<b>2 Príbuzné práce</b>	<b>13</b>
<b>3 Smerová štatistika</b>	<b>14</b>
3.1 Von Misesova distribúcia . . . . .	14
3.2 Zmes von Misesových distribúcií . . . . .	19
<b>4 Evolúcia</b>	<b>20</b>
4.1 Evolučné výpočty . . . . .	20
4.2 Evolučné algoritmy . . . . .	21
4.3 Evolučné stratégie . . . . .	22
4.4 Covariance Matrix Adaptation - Evolution Strategy . . . . .	23
<b>5 Metodológia a implementácia</b>	<b>33</b>
5.1 Metodológia von Misesovej distribúcie . . . . .	33
5.2 Metodológia algoritmu CMA-ES . . . . .	35
5.3 Implementácia algoritmu CMA-ES . . . . .	37
5.4 Implementácia programu . . . . .	40
5.5 Štatistické testy . . . . .	42
<b>6 Testovanie</b>	<b>44</b>
6.1 Testovanie CMA-ES algoritmu . . . . .	44
6.2 Testovanie výslednej aplikácie . . . . .	46
<b>7 Záver</b>	<b>60</b>
<b>Literatúra</b>	<b>61</b>
<b>Prílohy</b>	<b>65</b>
<b>A Obsah priloženého CD</b>	<b>65</b>

## Zoznam použitých skratiek a symbolov

BFGS	– Broyden Fletcher Goldfarb Shanno Algorithm
CMA	– Covariance Matrix Adaptation
CMA-ES	– Covariance Matrix Adaptation Evolution Strategy
CSA	– Cumulative Step-Size Adaptation
DE	– Differential Evolution
EA	– Evolution Algorithm
ES	– Evolution Strategy
GA	– Genetic Algorithm
NES	– Natural Evolution Strategy
NEWUOA	– NEW Unconstrained Optimization Algorithm
PDF	– Probability Density Function
PSO	– Particle Swarm Optimization
RMSD	– Root-Mean-Square Deviation
RMSE	– Root-Mean-Square Error



## Zoznam obrázkov

1	Funkcia hustoty pravdepodobnosti von Misesovej distribúcie [49]. . . . .	15
2	Distribučná funkcia von Misesovej distribúcie [49]. . . . .	15
3	Veterná ruža distribúcie pre rôzne $\kappa$ [14]. . . . .	16
4	Veterná ruža distribúcie pre rôzne $\mu$ [14]. . . . .	16
5	Ukážka priebehu CMA-ES algoritmu [50]. . . . .	24
6	Funkcia hustoty pravdepodobnosti viacrozmerného normálneho rozdelenia [51]. .	36
7	Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Drahany . . .	48
8	Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Drahany . . . . .	48
9	Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Kámen . . . .	49
10	Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Kámen . . . . .	50
11	Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Lipná . . . . .	51
12	Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Lipná . . . . .	51
13	Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Maletín . . . .	52
14	Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Maletín . . . . .	53
15	Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Veselí 26367 .	54
16	Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Veselí 26367 . . . . .	55
17	Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Veselí 26368 .	56
18	Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Veselí 26368 . . . . .	56
19	Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Mildred Lake .	58
20	Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Mildred Lake . . . . .	58

## Zoznam tabuliek

1	Počiatkové parametre jednotlivých von Misesových distribúcií - Drahnany . . . . .	47
2	Výsledky optimalizácie distribúcie smeru vetra pre Drahnany . . . . .	48
3	Počiatkové parametre jednotlivých von Misesových distribúcií - Kámen . . . . .	49
4	Výsledky optimalizácie distribúcie smeru vetra pre Kámen . . . . .	50
5	Počiatkové parametre jednotlivých von Misesových distribúcií - Lipná . . . . .	50
6	Výsledky optimalizácie distribúcie smeru vetra pre Lipná . . . . .	51
7	Počiatkové parametre jednotlivých von Misesových distribúcií - Maletín . . . . .	52
8	Výsledky optimalizácie distribúcie smeru vetra pre Maletín . . . . .	53
9	Počiatkové parametre jednotlivých von Misesových distribúcií - Veselí 26367 . . .	53
10	Výsledky optimalizácie distribúcie smeru vetra pre Veselí 26367 . . . . .	54
11	Počiatkové parametre jednotlivých von Misesových distribúcií - Veselí 26368 . . .	55
12	Výsledky optimalizácie distribúcie smeru vetra pre Veselí 26368 . . . . .	56
13	Porovnanie najlepšieho, priemerného a najhoršieho RMSE optimalizovanej distribúcie s RMSE analytického modelu pre každú lokáciu . . . . .	57
14	Počiatkové parametre jednotlivých von Misesových distribúcií - Mildred Lake . .	57
15	Výsledky optimalizácie distribúcie smeru vetra pre Mildred Lake . . . . .	58
16	Porovnanie najlepšieho, priemerného a najhoršieho $\chi^2$ optimalizovanej distribúcie s $\chi^2$ analytického modelu . . . . .	59

# 1 Úvod

Smerové dáta zohrávajú dôležitú úlohu v mnohých vedeckých a aplikovaných oboroch. Používajú sa na modelovanie širokého rozsahu fenoménov z rozličných oblastí, ako napríklad medicínske výskumy (syndróm náhlého úmrtia novorodencov [1]), v spracovaní obrazu pre spracovanie kamerových záznamov (detekcia abnormálneho správania pre policajný zbor [2]), predpoveď spotreby paliva [3] a rôzne aplikácie pre modelovanie vlastností vetra (identifikácia zdrojov znečistenia [10], ročné modelovanie smeru vetra [7] a posúdenie energetického potenciálu vetra [4]).

V posledných rokoch sa vietor stáva hlavným zdrojom čistej energie. Celosvetová ročná miera použitej veternej energie ukazuje, že na konci roka 2014 predstavovala inštalovaná kapacita veterných elektrární 369 GW. Oproti roku 2013 to predstavuje 16% nárast. Veterná energia sa podieľa približne 0,5% na svetovej výrobe elektriny [5].

Veterné elektrárne generujú elektrickú energiu pomocou veterných turbín. Veterná turbína konvertuje kinetickú energiu z vetra a vyrába elektrickú energiu. V dnešnej dobe sa veterné turbíny vyrábajú v širokom intervale od vertikálnych turbín až po horizontálne turbíny. Najmenšie turbíny nachádzajú využitie v karavanoch (napríklad, pre nabíjanie batérií) alebo tiež v doprave (napríklad, pre napájanie dopravného značenia). Väčšie turbíny nachádzajú použitie v domácnostiach ako záložný zdroj energie. Najväčšie turbíny sa zhromažďujú do skupín, ktoré označujeme ako veterné farmy a predstavujú dôležitý zdroj obnoviteľnej energie. Používajú sa v mnohých krajinách ako súčasť stratégie s cieľom znížiť závislosť na fosilných palivách.

Pri veterných turbínach rozlišujeme dva základné typy: s otáčajúcou sa a neotáčajúcou sa gondolou. Gondola je miesto, v ktorom sú ukryté generátor, prevodovka, prevody a brzdné mechanizmy turbíny. Otáčajúca sa gondola umožňuje natočenie turbíny podľa aktuálneho smeru vetra a robí z rýchlosti vetra jediný, najdôležitejší parameter ovplyvňujúci množstvo generovanej energie. U veternej turbíny s fixovanou gondolou sa stáva dôležitým parametrom naopak smer vetra [6].

Ako sme sa v predchádzajúcom odstavci dozvedeli, smer vetra a rýchlosť vetra sú najdôležitejšie premenné pre generovanie, integráciu a riadenie veternej energie. Presné modelovanie vlastností vetra je dôležité pre predpoklad množstva vygenerovanej veternej energie. Tieto predpoklady sú rozhodujúce pre efektívnu a bezpečnú prácu veterných turbín. Odhad množstva generovanej energie je preto nevyhnutnou súčasťou zelenej energie a získava význam v rozvinutých, ale aj rozvíjajúcich sa regiónoch. Analýza veterných vlastností pre špecifické miesto zohráva dôležitú úlohu pri vyberaní umiestnenia pre nové veterné zariadenia, a má tiež dopad na cenu vyrobenej elektrickej energie, ako aj na návrat investícií.

Aby sme mohli správne odhadovať potenciál veternej energie, musí byť navrhnutý zodpovedajúci štatistický model, ktorý analyzuje charakteristiky vetra [7]. Prvoradým cieľom analýzy charakteristík je identifikovať funkciu hustoty pravdepodobnosti (PDF) [8]. Avšak pre tento prípad nemôžno použiť diskkrétne modely, ktoré sa zvyčajne používajú pre analýzu smerovej rýchlosti vetra. Musia byť použité modely, založené iba na údajoch z histogramov. Tieto modely

rozdeľujú smerové dáta do oblastí alebo sektorov.

Druhý prípad kde môžeme použiť modelovanie smeru vetra je v oblasti prenosu a dodávky elektrickej energie. Výskumom v tejto oblasti sa snažíme o lepšie využitie prenosových liniek a nástrojov. Metóda dynamického tepelného výkonu odhaduje prenosovú kapacitu pomocou skutočných prevádzkových podmienok, a tak výrazne zlepšuje priepustnosť systému [9]. Pretože parametre vetra majú významný vplyv na tepelný výkon, je veľmi dôležité pre odhad skutočných prenosových limitov mať znalosť o funkcii hustoty pravdepodobnosti.

Táto práca popisuje metodológiu pre analýzu distribúcie smeru vetra. Táto analýza môže byť použitá na účely, ktoré sú popísané vyššie. Pre reprezentáciu veternej ruže je použitý pravdepodobnostný model spojitej premennej, založený na von Misesovej distribúcii. Parametre modelu sú stanovené pomocou evolučnej stratégie založenej na adaptácii kovariančnej matice. Inicializácia modelu je založená na rozdelení dát, ktoré sú navzájom od seba v rovnakej vzdialenosti. Kvalita vypočítanej distribúcie je potom štatisticky ohodnotená pomocou Pearsonovho  $\chi^2$ -testu dobrej zhody alebo pomocou RMSE testu.

Táto práca je organizovaná do nasledujúcich kapitol. Kapitola 2 popisuje príbuzné práce, ktoré sa zaoberali rovnakým problémom. V kapitole 3 nájdeme informácie o smerovej štatistike a teoretický popis von Misesovej distribúcie, zatiaľ čo kapitola 4 popisuje algoritmus CMA-ES. Kapitola 5 obsahuje popis toho, ako bola implementovaná výsledná aplikácia, a súčasne popisuje metodiku ako pomocou evolučných stratégií stanoviť zodpovedajúce štatistické rozdelenie. Výsledky testovania aplikácie sú popísané v kapitole 6. V kapitole 7 sú vyvedené hlavné závery tejto práce.

## 2 Príbuzné práce

Existuje veľa prístupov pre modelovanie a analýzu vlastností vetra. Väčšina z nich sa zameriava na krátkodobé alebo dlhodobé modelovanie rýchlosti vetra alebo na spojené modelovanie rýchlosti vetra a jeho smeru [7]. Existujú však aj práce, ktoré analyzujú, modelujú a odhadujú iba smer vetra [10]. Algoritmy použité pre modelovanie smeru vetra zahŕňajú detekciu lokálnych gradientov na obrázkoch zo syntetického svetelného radaru [11], vyhodnocovanie konečnej zmesi von Misesových distribúcií pomocou metódy najmenších štvorcov alebo pomocou minimalizačných algoritmov [7], prístupy založené na autoregresívnom pohyblivom priemere [12], a tiež niekoľko metaheuristických alebo biologicky inšpirovaných algoritmov [11].

Určovanie smeru vetra bolo súčasťou komplexných metód pre identifikáciu zdrojov znečistenia v práci Allena a kolektívu [10]. Hlavným cieľom výskumu bolo identifikovať povrchové smery vetra, a ďalšie detailné charakteristiky zdrojov znečistenia. Na zistenie zhody medzi parametrami modelu šírenia znečistenia a reálnymi dátami získanými pomocou senzorov, boli použité genetické algoritmy.

Použitie genetických systémov založených na fuzzy logike pre modelovanie lokálnych veterných podmienok bolo preskúmané de la Rosom a kolektívom [13]. Bol použitý fuzzy systém generovaný pomocou genetického algoritmu pre prispôbienie štatistického modelu veterných vlastností, aby sa lepšie zhodoval s miestnymi podmienkami na určitom mieste.

Heckenbergerová a kolektív [14], [15] použili jednoduchý genetický algoritmus pre nájdenie parametrov konečnej zmesi von Misesových distribúcií, aby modelovali ročnú pravdepodobnosť smeru vetra v špecifických lokáciách. Vo svojej práci pokračovali, a namiesto GA použili populárnu biologicky inšpirovanú metódu Particle Swarm Optimization. V oboch prípadoch dosiahli veľmi dobré výsledky.

Štyri modely rýchlosti a smeru vetra založené na adaptívnych neuro-fuzzy koncových systémoch a rôznych pravdepodobnostných rozdeleniach (Weibullovo, Frechetovo, Gumbelovo a spojené rozdelenie) boli popísané Shamshirbandom a kolektívom [16]. Štúdia ukázala veľmi dobré výsledky.

V tejto práci sa po dohode s vedúcim práce rozhodlo o použití pomerne novej biologicky inšpirovanej metóde Covariance Matrix Adaptation - Evolution Strategy. Tento algoritmus optimalizuje parametre konečnej zmesi von Misesových distribúcií, ktorá popisuje ročné rozloženie vetra v rôznych oblastiach. Tento algoritmus bol zvolený, pretože je designovaný pre optimalizáciu reálnych parametrov a nachádza uplatnenie v reálnom svete.

### 3 Smerová štatistika

Smerová štatistika je disciplína štatistiky, ktorá sa zaoberá smermi, osami a rotáciami v  $\mathcal{R}^n$ . Fakt, že 0 stupňov a 360 stupňov sú identické uhly (napríklad 180 stupňový uhol nie je strednou hodnotou 2 stupňov a 358 stupňov) nám poskytuje ilustráciu, že sú potrebné špeciálne štatistické metódy pre analýzu týchto dát. Ďalšími príkladmi dát, ktoré môžu byť považované za smerové dáta sú vyvíjajúce sa časové periódy, smery kompasu alebo uhly, orientácie a rotácie molekúl.

Dáta, ktoré popisujú smer sú veľmi dôležité v mnohých vedeckých a aplikovaných oblastiach. Tento typ dát, však nemôže byť analyzovaný klasickými štatistickými metódami, a preto musia byť použité špeciálne typy štatistických distribúcií pre modelovanie kruhových dát. V tejto práci bola pre modelovanie smerových dát použitá von Misesova distribúcia.

#### 3.1 Von Misesova distribúcia

V teórii pravdepodobnosti a smerovej štatistike, je von Misesova distribúcia spojená pravdepodobnostná distribúcia na kruhu. Je to blízka aproximácia obalenej normálnej distribúcie, ktorá je kruhovou analógiou normálnej distribúcie. Voľný rozptylový uhol  $\theta$  na kruhu je obalená, normálne distribuovaná náhodná veličina s neobaleným rozptylom, ktorý s časom rastie lineárne. Na druhej strane, von Misesova distribúcia je stacionárna distribúcia difúzneho procesu na kružnici s harmonickým potenciálom, teda inak povedané s preferovanou orientáciou [17]. Von Misesova distribúcia je distribúcia s maximálnou entropiou pre danú predpokladanú hodnotu  $z = e^{i\theta}$ .

Pravdepodobnostná funkcia hustoty von Misesovej distribúcie pre uhol  $x$  je daná predpisom [18]:

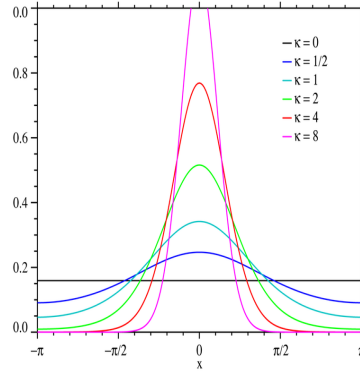
$$f(x|\mu, \kappa) = \frac{e^{\kappa \cos(x-\mu)}}{2\pi I_0(\kappa)}, \quad (1)$$

kde  $I_0(\kappa)$  je modifikovaná Besselova funkcia nultého radu:

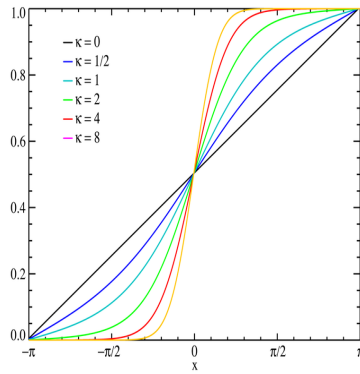
$$I_0(\kappa) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{\kappa \cos(\theta)} d\theta = \sum_{k=0}^{\infty} \frac{1}{(k!)^2} \left(\frac{\kappa}{2}\right)^{2k}. \quad (2)$$

Parametre  $\mu$  a  $1/\kappa$  sú analógiou ku  $\mu$  a  $\sigma^2$  (stredná hodnota a rozptyl) v normálnej distribúcii:

- $\mu$  je miera lokácie (distribúcia je zhľukovaná okolo  $\mu$ ) a
- $\kappa$  je miera koncentrácie (vzájomná miera disperzie, takže  $1/\kappa$  je analógia ku  $\sigma^2$ )
  - ak je  $\kappa$  nulová, distribúcia je uniformná, a pre malé  $\kappa$  je veľmi blízka uniformnosti,
  - ak je  $\kappa$  veľká, distribúcia sa stáva veľmi koncentrovaná okolo uhla  $\mu$  a  $\kappa$  sa stáva mierou koncentrácie. S rastúcou  $\kappa$  sa distribúcia stáva normálnou distribúciou so strednou hodnotou  $\mu$  a rozptylom  $1/\kappa$ .



Obr. 1: Funkcia hustoty pravdepodobnosti von Misesovej distribúcie [49].



Obr. 2: Distribučná funkcia von Misesovej distribúcie [49].

Hustota pravdepodobnosti môže byť vyjadrená ako séria Besselových funkcií:

$$f(x|\mu, \kappa) = \frac{1}{2\Pi} \left( 1 + \frac{2}{I_0(\kappa)} \sum_{j=1}^{\infty} I_j(\kappa) \cos[j(x - \mu)] \right), \quad (3)$$

kde  $I_j(x)$  je modifikovaná Besselová funkcia  $j$ -tého radu.

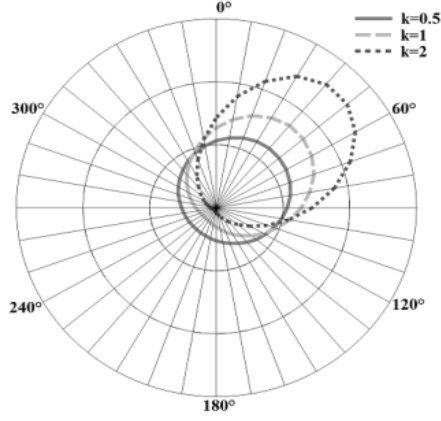
Kumulatívna distribučná funkcia sa nedá jednoducho analyticky vyjadriť, a preto najlepšie je ju nájsť integráciou predpisu 3. Neurčitý integrál hustoty pravdepodobnosti je [19]:

$$\Phi(x|\mu, \kappa) = \int f(t|\mu, \kappa) dt = \frac{1}{2\Pi} \left( x + \frac{2}{I_0(\kappa)} \sum_{j=1}^{\infty} I_j(\kappa) \sin[j(x - \mu)] \right). \quad (4)$$

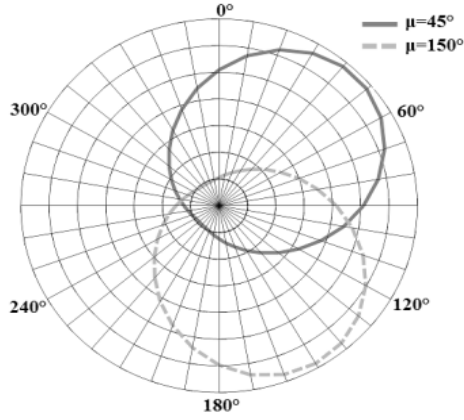
Kumulatívna distribučná funkcia má potom tvar [19]:

$$F(x|\mu, \kappa) = \Phi(x|\mu, \kappa) - \Phi(x_0|\mu, \kappa). \quad (5)$$





Obr. 3: Veterná ruža distribúcie pre rôzne  $\kappa$  [14].



Obr. 4: Veterná ruža distribúcie pre rôzne  $\mu$  [14].

### 3.1.1 Momenty

Momenty von Misesovej distribúcie sa zvyčajne vyratávajú ako momenty  $z = e^{i\theta}$  namiesto samotného uhlu  $x$ . Tieto momenty sa označujú ako kruhové momenty. Rozptyl vypočítaný z týchto momentov sa označuje ako kruhový rozptyl.

$n$ -tý moment  $z$  je [19]:

$$m_n = \langle z^n \rangle = \int_{\Gamma} z^n f(x|\mu, \kappa) dx = \frac{I_{|n|}(\kappa)}{I_0(\kappa)} e^{in\mu}, \quad (6)$$

kde integrál je cez ľubovoľný interval  $\Gamma$  dĺžky  $2\pi$ . Pri výpočte integrálu 6 používame fakt, že  $z^n = \cos(nx) + i \sin(nx)$  a Besselovskú funkciu:

$$I_n(\kappa) = \frac{1}{\pi} \int_0^\pi e^{\kappa \cos(x)} \cos(nx) dx. \quad (7)$$

Stredná hodnota  $z$  je potom [19]:

$$m_1 = \frac{I_1(\kappa)}{I_0(\kappa)} e^{i\mu} \quad (8)$$

a stredná hodnota  $x$  sa potom berie ako  $\mu$ . Rozptyl  $z$ , alebo kruhový rozptyl  $x$  je potom [19]:

$$\text{var}(x) = 1 - E[\cos(x - \mu)] = 1 - \frac{I_1(\kappa)}{I_0(\kappa)}. \quad (9)$$

### 3.1.2 Limity

Pre veľké  $\kappa$  hodnoty sa distribúcia stáva normálnou distribúciou [20]:

$$\lim_{\kappa \rightarrow \infty} f(x|\mu, \kappa) = \frac{1}{\sigma\sqrt{2\Pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right], \quad (10)$$

kde  $\sigma^2 = 1/k$ . Pre malé  $\kappa$  sa stáva uniformnou distribúciou [20]:

$$\lim_{\kappa \rightarrow 0} f(x|\mu, \kappa) = U(x), \quad (11)$$

kde interval uniformnej distribúcie  $U(x)$  je vybratý z intervalu dĺžky  $2\Pi$ .

### 3.1.3 Odhad parametrov

Séria  $N$  meraní  $z_n = e^{i\theta_n}$  získaných z von Misesovej distribúcie môže byť použitá pre odhad určitých parametrov distribúcie. Priemer série  $z$  je definovaný ako:

$$\bar{z} = \frac{1}{N} \sum_{n=1}^N z_n \quad (12)$$

a jeho predpokladaná hodnota bude iba prvý moment [20]:

$$\langle \bar{z} \rangle = \frac{I_1(\kappa)}{I_0(\kappa)} e^{i\mu}. \quad (13)$$

Ak berieme  $z_n$  ako množinu vektorov v komplexnej rovine,  $\bar{R}^2$  štatistika je druhá mocnina dĺžky priemerného vektora:

$$\bar{R}^2 = \bar{z}\bar{z}^* = \left(\frac{1}{N} \sum_{n=1}^N \cos \theta_n\right)^2 + \left(\frac{1}{N} \sum_{n=1}^N \sin \theta_n\right)^2 \quad (14)$$

a jeho očakávaná hodnota je [20]:

$$\langle \bar{R}^2 \rangle = \frac{1}{N} + \frac{N-1}{N} \frac{I_1(\kappa)^2}{I_0(\kappa)^2}. \quad (15)$$

Inými slovami, štatistika

$$R_e^2 = \frac{N}{N-1} \left( \bar{R}^2 - \frac{1}{N} \right) \quad (16)$$

bude odhad  $\frac{I_1(\kappa)^2}{I_0(\kappa)^2}$  a riešenie predpisu  $R_e = \frac{I_1(\kappa)}{I_0(\kappa)}$  pre  $\kappa$  prinesie lepší odhad pre  $\kappa$ .

### 3.1.4 Distribúcia strednej hodnoty

Distribúcia strednej hodnoty  $\bar{z} = \bar{R}e^{i\bar{\theta}}$  pre von Misesovu distribúciu je daná predpisom[21]:

$$P(\bar{R}, \bar{\theta}) d\bar{R} d\bar{\theta} = \frac{1}{(2\Pi I_0(\kappa))^N} \int_{\Gamma} \prod_{n=1}^N \left( e^{\kappa \cos(\theta_n - \mu)} d\theta_n \right) = \quad (17)$$

$$= \frac{e^{\kappa N \bar{R} \cos(\bar{\theta} - \mu)}}{I_0(\kappa)^N} \left( \frac{1}{(2\Pi)^N} \int_{\Gamma} \prod_{n=1}^N d\theta_n \right), \quad (18)$$

kde  $N$  je počet meraní a  $\Gamma$  pozostáva z intervalov dĺžky  $2\Pi$  v premenných,  $\bar{R}$  a  $\bar{\theta}$  sú konštanty, kde  $\bar{R}$  resultát strednej hodnoty [20]:

$$\bar{R}^2 = \|\bar{z}\|^2 = \left( \frac{1}{N} \sum_{n=1}^N \cos(\theta_n) \right)^2 + \left( \frac{1}{N} \sum_{n=1}^N \sin(\theta_n) \right)^2 \quad (19)$$

a  $\bar{\theta}$  je uhol strednej hodnoty:  $\bar{\theta} = \arg(\bar{z})$ .

### 3.1.5 Entropia

Informačná entropia von Misesovej distribúcie je definovaná ako [20]:

$$H = - \int_{\Gamma} f(\theta, \mu, \kappa) \ln(f(\theta, \mu, \kappa)) d\theta, \quad (20)$$

kde  $\Gamma$  je ľubovoľný interval dĺžky  $2\Pi$ . Logaritmus hustoty von Misesovej distribúcie je:

$$\ln(f(\theta, \mu, \kappa)) = -\ln(2\Pi I_0(\kappa)) + \kappa \cos(\theta). \quad (21)$$

Charakteristická funkčná reprezentácia von Misesovej distribúcie je:

$$f(\theta, \mu, \kappa) = \frac{1}{2\Pi} \left( 1 + 2 \sum_{n=1}^{\infty} \phi_n \cos(n\theta) \right), \quad (22)$$

kde  $\phi_n = I_{\|n\|}(\kappa)/I_0(\kappa)$ . Entropia môže byť potom zapísaná aj ako:

$$H = \ln(2\Pi I_0(\kappa)) - \kappa \phi_1 = \ln(2\Pi I_0(\kappa)) - \kappa \frac{I_1(\kappa)}{I_0(\kappa)}. \quad (23)$$

Pre  $\kappa = 0$  sa von Misesova distribúcia stáva kruhovou uniformnou distribúciou.

### 3.2 Zmes von Misesových distribúcií

Jednoduchá von Misesova distribúcia je unimodálna a symetrická, čo znamená, že môže aproximovať iba smerové dáta s jedným prevládajúcim smerom. Keď dáta obsahujú viac ako jeden prevládajúci smer je potrebné použiť zmes týchto distribúcií. Model, ktorý vznikne zmiešaním von Misesových distribúcií je daný súčtom jednoduchých von Misesových funkcií hustoty pravdepodobnosti nasledovne:

$$\phi(\theta, \omega_j, \mu_j, \kappa_j) = \sum_{j=1}^N \omega_j f_j(\theta, \mu_j, \kappa_j), \quad (24)$$

kde  $N$  je počet hlavných smerov,  $j$  je index jednotlivých jednoduchých von Misesových funkcií hustoty pravdepodobnosti s parametrami  $\mu_j$ ,  $\kappa_j$  a  $\omega_j$  je nezáporná váha spĺňajúca nasledovné:

$$0 \leq \omega_j \leq 1 \quad (25)$$

$$\sum_{j=1}^N \omega_j = 1. \quad (26)$$

Premenné  $\mu_j$ ,  $\kappa_j$  a  $\omega_j$  spolu tvoria vektor  $\mathbf{v}$ :

$$\mathbf{v} = (\mu_j, \kappa_j, \omega_j) = (\mu_1, \dots, \mu_k, \kappa_1, \dots, \kappa_k, \omega_1, \dots, \omega_k). \quad (27)$$

Pravdepodobnosť, že smer vetra je vnútri nejakého určitého uhlového sektora, môže byť dosiahnutá integráciou 24 medzi dvoma sektorovými hraničnými hodnotami  $\theta$ .

$$P(\theta \in \langle \theta_l, \theta_u \rangle) = \int_{\theta_l}^{\theta_u} \phi(\theta, \mathbf{v}) d\theta. \quad (28)$$

Kumulatívna distribučná funkcia zmesi von Misesových distribúcií je potom:

$$F(x|\mu, \kappa) = \sum_{j=1}^N \frac{\omega_j}{2\pi I_0(\kappa_j)} \int_0^\theta e^{\kappa_j \cos(\theta - \mu_j)} d\theta. \quad (29)$$

Predpis 29 musí byť vypočítaný numericky. Narozdiel od jednoduchej von Misesovej distribúcie sa parametre zmiešaného modelu nedajú odhadnúť analyticky, a preto musia byť použité numerické optimalizačné metódy.

## 4 Evolúcia

Evolúcia je proces rastu a rozvoja, ktorý v sebe nesie zmenu. Slovo pochádza z latinského slova *evolutio*, čo znamená rozvinutie, a pred koncom 19. storočia sa tento pojem výhradne používal na cielené, naprogramované procesy ako napríklad embryonálny vývoj. Po vydaní publikácie Charlesa Darwina *O pôvode druhov* v roku 1859 sa evolúcia priamo spojila s biologickou evolúciou. Biologická evolúcia predstavuje necielený proces pre vývoj rôznorodých foriem života zo spoločného predka. Jednou z foriem evolúcie sú evolučné výpočty, ktoré sa pokúšajú napodobniť procesy biologickej evolúcie pomocou počítačového programu najčastejšie ako optimalizačné techniky.

### 4.1 Evolučné výpočty

V informatike sú evolučné výpočty oblasťou umelej inteligencie, ktoré môžu byť definované typom algoritmov, ktoré obsahujú. Tieto algoritmy, nazývané evolučné algoritmy, sú založené na adaptácii Darwinových evolučných princípov. Technicky patria do skupiny algoritmov riešiacich problémy metódou “pokus omyl” a možno ich považovať za globálne optimalizačné metódy s metaheuristickým alebo stochastickým optimalizačným charakterom. Vyznačujú sa tiež použitím populácie kandidátnych riešení. Najčastejšie sa používajú na riešenie “black box” problémov, často v kontexte s vysokou cenou optimalizácie. Evolučné výpočty používajú iteratívny postup, ako je napríklad rast alebo vývoj populácie. Táto populácia je potom vybratá v riadenom náhodom prehľadávaní použitím paralelného spracovania, aby sa zachoval požadovaný koniec.

Použitie Darwinových princípov pre automatizované riešenie problémov pochádza z 50. rokov 20. storočia. V 60. rokoch sa začali vyvíjať tri nezávislé interpretácie tohto nápadu na troch rôznych miestach. Evolučné programovanie bolo predstavené Lawrenceom J. Fogelom [22] v USA, pričom John Henry Holland [23] nazval túto metódu genetické algoritmy. V Nemecku Ingo Rechenberg [24] a Hans-Paul Schwefel [25] predstavili evolučné stratégie. Tieto oblasti sa vyvíjali nezávisle na sebe približne 15 rokov. Od začiatku 90. rokov sú zjednotené ako rôzne reprezentácie tejto istej technológie, ktoré označujeme evolučné výpočty. V rovnakom období sa tiež objavila štvrtá skupina - genetické programovanie [26].

Simulácie evolúcie použitím evolučných algoritmov a umelého života začali prácou Nilsa Aalla Barricelliho [27] v 60. rokoch, a boli rozšírené Alexom Fraserom, ktorý publikoval sériu článkov o umelej selekcii [28]. Umelá selekcia sa stala rozšírenou optimalizačnou metódou k čomu prispela práca Inga Rechenberga v 70. rokoch, ktorý použil evolučné stratégie k riešeniu strojarských problémov [24]. Genetické algoritmy sa stali populárne vďaka práci Johna Hollanda [23]. Nárast akademického záujmu umožnil praktické aplikácie, napríklad evolúcia počítačových programov [29]. Evolučné algoritmy sa dnes používajú na riešenie multi-dimenzionálnych prob-

lémov, pretože ich riešia efektívnejšie ako programy vytvorené ľudskou rukou a poskytujú lepšiu optimalizáciu [30].

## 4.2 Evolučné algoritmy

Evolučné algoritmy, ktoré sú podmnožinou evolučných výpočtov, sú populačne založené metaheuristické optimalizačné algoritmy. Evolučné algoritmy používajú mechanizmy inšpirované biologickou evolúciou, ako napríklad reprodukcia, mutácia, rekombinácia a selekcia. Jednotlivé kandidátne riešenia optimalizačného problému hrajú úlohu jedincov populácie a fitness funkcia (funkcia vhodnosti) určuje kvalitu riešenia. Evolúcia populácie potom prebieha po opakovanom použití vyššie uvedených operácií. Evolučné algoritmy často vykonávajú dobrú aproximáciu riešení pre všetky typy problémov, pretože nevytvárajú žiadne predpoklady o fitness landscape (oblasť vhodnosti). Dôsledkom toho je, že tieto algoritmy sú úspešné v rôznych oblastiach, ako strojárstvo, umenie, biológia, ekonomika, marketing, genetika, operačný výskum, robotika, sociálne vedy, fyzika, politika a chémia.

Vo väčšine reálnych aplikácií evolučných algoritmov je výpočetná zložitosť obmedzujúcim faktorom. Hlavným dôvodom zložitosti je ohodnotenie fitness funkcie. Aproximácia fitness funkcie je jedným z možných riešení tohto problému. Na druhej strane však jednoduchý evolučný algoritmus môže vyriešiť zložitý problém, preto nemôže existovať priama súvislosť medzi zložitou algoritmu a zložitou problémom.

### 4.2.1 Implementácia biologických procesov

1. Generovanie náhodnej počiatočnej populácie jedincov
2. Vyhodnotenie fitness pre každého jedinca populácie
3. Opakované pracovanie na populácii, kým nedôjde k terminácii
  - (a) Výber jedincov s najlepšou fitness pre reprodukciu - rodičia
  - (b) Vytvorenie nových jedincov pomocou kríženia a mutácie - potomstvo
  - (c) Vyhodnotenie fitness nových jedincov
  - (d) Nahradenie jedincov s najhoršou fitness novými jedincami

### 4.2.2 Typy evolučných algoritmov

- **Genetické algoritmy** - Najpopulárnejší typ evolučných algoritmov. Hľadajú riešenie problému, ktoré je vo forme reťazcov čísel (tradične binárnych), aplikáciou rekombinácie a mutácie.
- **Genetické programovanie** - Riešenie je vo forme počítačových programov a ich fitness je určené schopnosťou vyriešiť problém.

- **Evolučné programovanie** - Podobné genetickému programovaniu, ale štruktúra programu sa nemôže meniť a môžu sa meniť iba parametre.
- **Programovanie genetických výrazov** - Podobné genetickému programovaniu, ale spoločne s vývojom počítačových programov skúma možnosti ako zakódovať programy rôznych dĺžok do štruktúr s rovnakou dĺžkou.
- **Evolučné stratégie** - Pracujú s vektormi reálnych čísel, ktoré reprezentujú riešenia.
- **Diferenciálna evolúcia** - Taktiež pracuje s vektormi a je určená pre numerické optimalizačné metódy.
- **Neuroevolúcia** - Podobná genetickému programovaniu ale slúži na evolúcie neuronových sietí.

### 4.3 Evolučné stratégie

Evolučná stratégia je optimalizačná technika založená na nápadoch adaptácie a evolúcie. Evolučné stratégie používajú prirodzené problémovo-závislé reprezentácie, a ako primárne vyhľadávacie operátory mutáciu a selekciu. Vo všeobecnosti sa operátory používajú v cykle. Jedná iterácia cyklu sa nazýva generácia. Sekvencia generácií pokračuje, kým nie sú splnené terminačné kritériá. Pokiaľ ide o priestory prehľadávania, ktoré sú tvorené reálnymi číslami, mutácia sa vykonáva zvyčajne pridaním normálne distribuovanej náhodnej veličiny pre každú zložku vektora. Veľkosť kroku alebo sila mutácie je často riadená pomocou samo-adaptácie. Jednotlivé veľkosti krokov pre každú koordináciu alebo koleráciu medzi koordinátami sú buď riadené pomocou samo-adaptácie ale pomocou adaptácie kovariančnej matice (CMA-ES).

Selekcia v evolučných stratégiách je deterministická, a je založená iba na porovnávaní fitness hodnôt. Výsledný algoritmus je nemenný vzhľadom k monotónej transformácii objektívnej funkcie. Najjednoduchšia evolučná stratégia pracuje na populácii veľkosti 2: jeden bod (rodič) a druhý bod ako výsledok jeho mutácie. Iba ak je fitness hodnota potomka taká dobrá, ako je fitness hodnota rodiča, potomok sa stane rodičom v ďalšej generácii. V opačnom prípade sa na potomkov neprihliada. Toto je  $(1 + 1)$ -ES. Vo všeobecnosti,  $\lambda$  možných potomkov môže byť generovaných, aby súťažili s rodičmi a takýto typ sa nazýva  $(1 + \lambda)$ -ES. V tejto ES sa potomok s najlepšou fitness hodnotou stane rodičom v novej generácii, zatiaľ čo na aktuálneho rodiča sa neprihliada. Súčasne derivácie evolučných stratégií často používajú populáciu  $\mu$  rodičov, a tiež rekombináciu ako ďalší operátor. Tento typ sa nazýva  $(\mu/\rho +, \lambda)$ -ES a je menej náchylný na uviaznutie v lokálnom optime.



## 4.4 Covariance Matrix Adaptation - Evolution Strategy

Covariance Matrix Adaptation - Evolution Strategy je stochastická, bezderivačná metóda pre matematickú optimalizáciu nelineárnych alebo nekonvexných spojitých optimalizačných problémov [31]. Patrí do triedy evolučných algoritmov a evolučných výpočtov. Je založená na princípoch biologickej evolúcie, menovite opakovaním súhrnu variácií (prostredníctvom rekombinácie a mutácie), a selekcií sa v každej generácii (iterácii) generujú noví jedinci (kandidátne riešenia, označené ako  $\mathbf{x}$ ) pomocou variácií z aktuálnych rodičovských jedincov zvyčajne stochastickým spôsobom. V ďalšom kroku sú niektorí jedinci vybratí, aby sa stali rodičmi v ďalšej generácii. Výber je založený na fitness hodnote alebo objektívnej funkčnej hodnote  $f(\mathbf{x})$ . Týmto spôsobom sú v každej generácii generovaní jedinci s lepšími a lepšími  $f$ -hodnotami.

V evolučnej stratégii sa nové kandidátne riešenia vyberajú vzhľadom na viacrozmerné normálne distribúcie v  $\mathbb{R}^n$ . Rekombinácia sa viaže na vybratie novej strednej hodnoty distribúcie. Mutácia spôsobuje pridanie náhodného vektora s odchýlkou s nulovou strednou hodnotou. Párové závislosti medzi premennými v distribúcii sú reprezentované kovariančnou maticou. Adaptácia kovariančnej matice (CMA) je metóda úpravy kovariančnej matice tejto distribúcie [31]. Toto je užitočné najmä v prípade, že funkcia  $f$  je podmienená. Adaptácia kovariančnej matice sa viaže na učenie modelu druhého radu z podkladovej objektívnej funkcie. V porovnaní s klasickými metódami je využívaných menej predpokladov podkladovej objektívnej funkcie. Na učenie je využívané iba poradie medzi kandidátskymi riešeniami a nie derivácie alebo dokonca funkčné hodnoty.

### 4.4.1 Princípy

V CMA-ES algoritme sú využívané dva hlavné princípy pre adaptáciu parametrov prehľadávanej distribúcie.

**4.4.1.1 Prvý princíp** Tento princíp sa označuje ako metóda maximálnej vierohodnosti. Je založený na tom, že zvyšuje pravdepodobnosť úspešných kandidátnych riešení a vyhľadávacích krokov. Stredná hodnota distribúcie sa aktualizuje tak, aby vierohodnosť predošlého úspešne nájdeného kandidátneho riešenia bola čo najväčšia. Kovariančná matica distribúcie je aktualizovaná tak, aby vierohodnosť predchádzajúceho úspešného vyhľadávacieho kroku bola zvýšená. Obe aktualizácie môžeme chápať ako vzostup prirodzeného gradientu. Aj v dôsledku toho, vykonáva CMA v každej iterácii analýzu základných zložiek úspešných vyhľadávacích krokov, pričom zachováva všetky hlavné osi [32].

**4.4.1.2 Druhý princíp** Sú zaznamenávané dve cesty časového priebehu strednej hodnoty distribúcie, nazývané tiež evolučné cesty. Tieto cesty obsahujú dôležité informácie o korelácii medzi krokmi idúcimi po sebe. Konkrétne, ak dva po sebe idúce kroky sú v rovnakom smere,

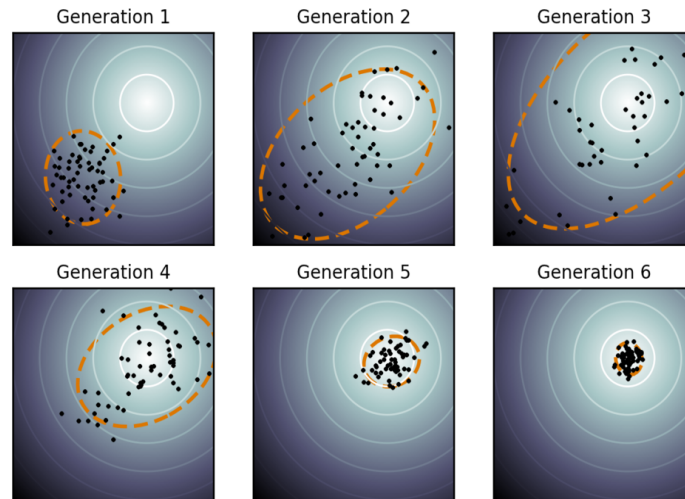
evolučné cesty sa predlžia. Evolučné cesty sa využívajú dvoma spôsobmi. Jeden spôsob používa procedúru adaptácie kovariančnej matice v mieste jedného úspešného vyhľadávacieho kroku, a uľahčuje, prípadne oveľa zrýchľuje nárast priaznivých smerov. Druhý spôsob sa používa na ďalšiu kontrolu veľkosti kroku. Táto kontrola veľkosti kroku sa zameriava na to, aby po sebe idúce kroky strednej hodnoty distribúcie boli ortogonálne. Táto kontrola efektívne zabráňuje predčasnej aproximácii, pričom dovoľuje rýchlu konvergenciu k optimu [32].

#### 4.4.2 Algoritmus

Najznámejším typom CMA-ES je  $(\mu/\mu_w, \lambda)$ -CMA-ES. V každom iteračnom kroku sa používa vážená kombinácia  $\mu$  najlepších z  $\lambda$  nových kandidátnych riešení na aktualizáciu parametrov distribúcie. Hlavný cyklus pozostáva z troch hlavných častí [31]:

- vzorkovanie nových riešení,
- preusporiadanie ovzorkovaných riešení podľa ich fitness hodnoty,
- aktualizácia vnútorných premenných na základe preusporiadaných riešení.

Nasledujúca ilustrácia 5 zobrazuje aktuálny optimalizačný priebeh s adaptáciou kovariančnej matice na jednoduchom dvojdimenzionálnom probléme. Sférická optimalizácia funkcie je zobrazená pevnými čiarami, ktoré zobrazujú miesta s rovnakou funkčnou hodnotou. Populácia (bodky) je oveľa väčšia ako je potrebné, ale jasne ukazuje ako sa distribúcia populácie (čiarkovaná čiara) mení v priebehu optimalizácie. Na tomto jednoduchom probléme sa populácia koncentruje okolo globálneho optima v priebehu pár generácií.



Obr. 5: Ukážka priebehu CMA-ES algoritmu [50].

#### 4.4.2.1 Pseudokód Pseudokód algoritmu vyzerá nasledovne:

Výber hodnoty  $\lambda$  (počet vzoriek počas iterácie, najmenej 2, všeobecne  $> 4$ );  
 Inicializácia  $m$ ,  $\sigma$ ,  $C = I$ ,  $p_\sigma = 0$ ,  $p_c = 0$  (stavové premenné);  
**while** *nie sú splnené terminačné kritériá* **do**  
     **for**  $i \leftarrow 1 \dots \lambda$  **do**  
          $x_i = \text{Vzorkovanie podľa normálneho rozdelenia (stredná hodnota} = m,$   
          $\text{rozptyl} = \sigma^2 C);$   
          $f_i = \text{Fitness}(x_i);$   
     **end**  
      $x_{1\dots\lambda} \leftarrow x_{s(1)\dots s(\lambda)} \mid s(i) = \text{Usporiadaj}(f_{1\dots \lambda}, i)$  (usporiadanie riešení podľa fitness);  
      $m' = m;$   
      $m \leftarrow \text{Aktualizuj } m(x_1, \dots, x_\lambda)$  (presun strednej hodnoty na lepšie riešenia);  
      $p_\sigma \leftarrow \text{Aktualizuj } p_\sigma(p_\sigma, \sigma^{-1} C^{-1/2}(m - m'))$  (aktualizácia izotropnej evolučnej cesty);  
      $p_c \leftarrow \text{Aktualizuj } p_c(p_c, \sigma^{-1}(m - m'), \|p_\sigma\|)$  (aktualizácia antiizotropnej evolučnej cesty);  
      $C \leftarrow \text{Aktualizuj } C(C, p_c, (x_1 - m')/\sigma, \dots, (x_\lambda - m')/\sigma)$  (aktualizácia kovariančnej matice);  
      $\sigma \leftarrow \text{Aktualizuj } \sigma(\sigma, \|p_\sigma\|)$  (aktualizácia veľkosti kroku pomocou izotropnej dĺžky cesty);  
**end**  
 Vráť  $m$  alebo  $x_1$ ;

Poradie piatich aktualizáčnych úloh je relevantné. V nasledujúcom odstavci je špecifikovaných päť aktualizáčnych rovníc pre päť stavových premenných.

Dané sú dimenzia prehľadávaného priestoru  $n$  a iteračný krok  $k$ . Päť stavových premenných je [33]:

- $m_k \in \mathbb{R}^n$ , stredná hodnota distribúcie a aktuálne najlepšie riešenie optimalizačného problému,
- $\delta_k > 0$ , veľkosť kroku,
- $C_k$ , symetrická a pozitívne definitná  $n \times n$  kovariančná matica s  $C_0 = I$  a
- $p_\delta \in \mathbb{R}^n$ ,  $p_c \in \mathbb{R}^n$ , dve evolučné cesty, inicializované ako nulový vektor.

Iterácia začína s  $\lambda > 1$  kandidátnych riešení  $x_i \in \mathbb{R}^n$  z viacrozmerného normálneho rozdelenia  $\mathcal{N}(m_k, \sigma_k^2 C_k)$  to je pre  $i = 1, \dots, \lambda$

$$x_i \sim \mathcal{N}(m_k, \sigma_k^2 C_k), \quad (30)$$

$$x_i \sim m_k + \sigma_k \times \mathcal{N}(0, C_k), \quad (31)$$

kde:

- $\sim$  - označuje, že na ľavej aj pravej strane je rovnaká distribúcia,
- $\mathcal{N}(0, C_k)$  - viacrozmerná normálna distribúcia s nulovou strednou hodnotou a kovariančnou maticou  $\mathbf{C}$ ,
- $x_i$  -  $i$ -tý jedinec v populácii,
- $m_k$  - stredná hodnota distribúcie,
- $\sigma_k$  - štandardná odchýlka distribúcie,
- $C_k$  - kovariančná matica,
- $k$  - číslo generácie.

Rovnica 31 ukazuje interpretáciu ako narušenie (mutáciu) aktuálne najlepšieho vektora  $m_k$  (vektor strednej hodnoty rozdelenia). Kandidátne riešenia  $x_i$  sú vypočítané ako objektívna funkcia  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  tak, aby boli minimalizované. Usporiadané kandidátne riešenia sa označia [33]:

$$\{x_{i:\lambda} \mid i = 1 \dots \lambda\} = \{x_i \mid i = 1 \dots \lambda\}, \quad (32)$$

$$f(x_{1:\lambda}) \leq \dots \leq f(x_{\mu:\lambda}) \leq f(x_{\mu+1:\lambda}) \dots, \quad (33)$$

nová stredná hodnota sa vyráta ako [33]:

$$m_{k+1} = \sum_{i=1}^{\mu} w_i x_{i:\lambda}, \quad (34)$$

$$m_{k+1} = m_k + \sum_{i=1}^{\mu} w_i (x_{i:\lambda} - m_k), \quad (35)$$

kde sa súčet pozitívnych váh  $w_1 \geq w_2 \geq \dots \geq w_{\mu} > 0$  rovná 1. Typicky,  $\mu \leq \lambda/2$  a váhy sú vybraté napríklad takto:  $\mu_w := 1 / \sum_{i=1}^{\mu} w_i^2 \approx \lambda/4$ . Jediná spätná väzba použitá z objektívnej funkcie je usporiadanie kandidátnych riešení vzhľadom k indexom  $i : \lambda$ . Veľkosť kroku  $\sigma_k$  sa aktualizuje pomocou *kumulatívnej adaptácie veľkosti kroku* (CSA), niekedy sa tiež označuje ako *kontrola dĺžky kroku*. Evolučná cesta  $p_{\sigma}$  sa aktualizuje najskôr [33]:

$$p_{\sigma} \leftarrow (1 - c_{\sigma})p_{\sigma} + \sqrt{1 - (1 - c_{\sigma})^2} \sqrt{\mu_w} C_k^{-1/2} \frac{m_{k+1} - m_k}{\sigma_k}, \quad (36)$$

$$\sigma_{k+1} = \sigma_k \times \exp\left(\frac{c_\sigma}{d_\sigma} \left(\frac{\|p_\sigma\|}{E \|\mathcal{N}(0, I)\|} - 1\right)\right), \quad (37)$$

kde:

- $c_\sigma^{-1} \approx n/3$  je spätný časový horizont pre evolučnú cestu  $p_\sigma$  a je väčší ako 1,
- $\mu_w = \left(\sum_{i=1}^{\mu} w_i^2\right)^{-1}$  je rozptylová efektívna selekčná váha a  $1 \leq \sigma_w \leq \sigma$  z definície  $w_i$ ,
- $C_k^{-1/2} = \sqrt{C_k}^{-1} = \sqrt{C_k^{-1}}$  je unikátna symetrická odmocnina inverznej matice k matici  $C_k$  a
- $d_\sigma$  je tlmiaci parameter zvyčajne blízko 1. Pre  $d_\sigma = \infty$  alebo  $c_\sigma = 0$  veľkosť kroku ostáva nezmenená.

Veľkosť kroku  $\sigma_k$  sa zvyšuje len vtedy ak  $\|p_\sigma\|$  je väčšia ako predpokladaná hodnota

$$E \|\mathcal{N}(0, I)\| = \sqrt{2}\Gamma((n+1)/2)/\Gamma(n/2), \quad (38)$$

$$E \|\mathcal{N}(0, I)\| \approx \sqrt{n}(1 - 1/(4n) + 1/(21n^2)), \quad (39)$$

a znižuje ak je menšia. Pre tento dôvod aktualizácia veľkosti kroku vedie k tomu, že po sebe idúce kroky  $C_{k-1}$  konjugujú, potom ako bola adaptácia úspešná [34].

$$\left(\frac{m_{k+2} - m_{k+1}}{\sigma_{k+1}}\right)^T C_k^{-1} \frac{m_{k+1} - m_k}{\sigma_k} \approx 0. \quad (40)$$

Nakoniec sa aktualizuje kovariančná matica, ale najskôr sa opäť aktualizuje evolučná cesta [33]:

$$p_c \leftarrow (1 - c_c)p_c + 1_{[0, \alpha\sqrt{n}]}(\|p_\sigma\|) \sqrt{1 - (1 - c_c)^2} \sqrt{\mu_w} \frac{m_{k+1} - m_k}{\sigma_k}, \quad (41)$$

$$C_{k+1} = (1 - c_1 - c_\mu + c_s)C_k + c_1 p_c p_c^T + c_\mu \sum_{i=1}^{\mu} w_i \frac{x_{i:\lambda} - m_k}{\sigma_k} \left(\frac{x_{i:\lambda} - m_k}{\sigma_k}\right)^T, \quad (42)$$

kde  $T$  znamená transpozíciu a

- $c_c^{-1} \approx n/4$  je spätný časový horizont pre evolučnú cestu  $p_c$  a je väčší ako 1,
- $\alpha \approx 1.5$  a charakteristická funkcia  $1_{[0, \alpha\sqrt{n}]}(\|p_\sigma\|)$  sa blíži 1 len vtedy ak  $\|p_\sigma\| \in [0, \alpha\sqrt{n}]$  alebo, inými slovami  $\|p_\sigma\| \leq \alpha\sqrt{n}$ , čo sa väčšinou stáva,
- $c_s = (1 - 1_{[0, \alpha\sqrt{n}]}(\|p_\sigma\|)^2)c_1 c_c (2 - c_c)$  robí čiastočné zmeny pre malé straty rozptylu v prípade, že indikátor je nulový,

- $c_1 \approx 2/n^2$  je rýchlosť učenia pre rank-one aktualizáciu kovariančnej matice a
- $c_\mu \approx \mu_w/n^2$  je rýchlosť učenia pre rank- $\mu$  aktualizáciu kovariančnej matice a nesmie prejsiahnuť  $1 - c_1$ .

Aktualizácia kovariančnej matice vedie k zvyšujúcej sa vierohodnosti pre  $p_c$  a pre  $(x_{i:\lambda} - m_k)/\sigma_k$  aby boli ovzorkované z  $\mathcal{N}(0, C_{k+1})$ . Toto tvorí jeden iteračný krok.

Počet kandidátnych vzoriek v iterácii  $\lambda$  nie je deterministický a môže sa meniť v širokom intervale. Malé hodnoty, ako napríklad  $\lambda = 10$ , vedú viac k tomu, že algoritmus vyhľadáva lokálne. Väčšie hodnoty, ako napríklad  $\lambda = 10n$  s defaultnou hodnotou  $\mu_w \approx \lambda/4$  robia vyhľadávanie viac globálne. Niekedy je algoritmus opakovane znovu spúšťaný so zvyšujúcim sa  $\lambda$  o 2 s každým novým spustením [35]. Okrem nastavenia  $\lambda$  (alebo  $\mu$  namiesto toho, ak je napríklad  $\lambda$  limitujúce počtom dostupných procesorov) ostatné vyššie zmienené parametre nie sú špecifikované vzhľadom na objektívnu funkciu, a preto ich používateľ nemusí nastavovať.

#### 4.4.3 Teoretické základy

Vzhľadom k parametrom distribúcie - stredná hodnota, rozptyl a kovariancia - normálne rozdelenie pravdepodobnosti pre vzorkovanie nových kandidátnych riešení je rozdelenie pravdepodobnosti s maximálnou entropiou v  $\mathbb{R}^n$ , čo znamená, že pre vzorkovanie rozdelenia sa používa distribúcia s minimálnym počtom predbežných informácií. V nasledujúcej časti sú popísané úvahy o aktualizácii rovníc CMA-ES.

**4.4.3.1 Variabilná metrika** CMA-ES implementuje stochastickú variabilnú metriku. Vo veľmi osobitnom prípade konvexnej kvadratickej objektívnej funkcie:

$$f(x) = \frac{1}{2}(x - x^*)^T H (x - x^*), \quad (43)$$

kovariančná matica  $C_k$  sa adaptuje na inverznú Hessovu maticu  $H$ , s malými náhodnými fluktuáciami. Vo všeobecnosti, na funkcii  $g \circ f$ , kde  $g$  je striktne rastúca a  $f$  je konvexne-kvadratická sa kovariančná matica  $C_k$  adaptuje na  $H^{-1}$  s malými náhodnými fluktuáciami.

**4.4.3.2 Aktualizácia s maximálnou vierohodnosťou** Aktualizačné rovnice pre strednú hodnotu a kovariančnú maticu zvyšujú vierohodnosť. Aktualizácia vektora strednej hodnoty  $m$  zvyšuje vierohodnosť v logaritmickej mierke:

$$m_{k+1} = \arg \max_m \sum_{i=1}^{\mu} w_i \log p_{\mathcal{N}}(x_{1:\lambda} | m), \quad (44)$$

kde

$$\log p_{\mathcal{N}}(x) = -\frac{1}{2} \log \det(2\pi C) - \frac{1}{2}(x - m)^T C^{-1}(x - m) \quad (45)$$

označuje logaritmicú vierohodnosť  $x$  z normálneho rozdelenia so strednou hodnotou  $m$  a hocijaku pozitívne definitnou maticou  $C$ . Môžeme vidieť, že  $m_{k+1}$  je nezávislé na matici  $C$ .

Rank- $\mu$  aktualizácia kovariančnej matice, čo znamená, že sčítanec, ktorý je v aktualizáčnej rovnici  $C_k$  najviac na pravej strane, zvyšuje vierohodnosť takto:

$$\sum_{i=1}^{\mu} w_i \frac{x_{1:\lambda} - m_k}{\sigma_k} \left( \frac{x_{1:\lambda} - m_k}{\sigma_k} \right)^T = \arg \max_C \sum_{i=1}^{\mu} w_i \log p_{\mathcal{N}} \left( \frac{x_{1:\lambda} - m_k}{\sigma_k} \middle| C \right) \quad (46)$$

pre  $\mu \geq n$ .  $p_{\mathcal{N}}(x|C)$  označuje vierohodnosť  $x$  z normálneho rozdelenia s nulovou strednou hodnotou a kovariančnou maticou  $C$ .

#### 4.4.3.3 Klesanie prirodzeného gradientu v priestore distribúcie Akimoto [36]

a Glasmachers [37] zistili nezávisle na sebe, že aktualizácia parametrov rozdelenia sa podobá klesaniu v smere prirodzeného gradientu očakávanej objektívnej funkcie  $Ef(x)$ . S nastavením parametrov  $c_{\sigma} = 0$  a  $c_1 = 0$ , teda bez kontroly veľkosti kroku a rank-one aktualizácie, sa CMA-ES algoritmus podobá na inštanciu prirodzených evolučných stratégií (NES). Prirodzený gradient je nezávislý na parametrizácii distribúcie. S ohľadom na parameter  $\theta$  distribúcie  $p$ , gradient  $Ef(x)$  môže byť vyjadrený ako:

$$\nabla_{\theta} E(f(x)|\theta) = \nabla_{\theta} \int_{\mathbb{R}^n} f(x) p(x) dx = E(f(x) \nabla_{\theta} \ln p(x|\theta)), \quad (47)$$

kde  $p(x) = p(x|\theta)$  závisí na parametrovom vektore  $\theta$ , tzv. skóre funkcii,  $\nabla_{\theta} \ln p(x|\theta) = \frac{\nabla_{\theta} p(x)}{p(x)}$ , čo indikuje relatívnu senzitivitu  $p$ , a preto je odhad braný s ohľadom na distribúciu  $p$ . Prirodzený gradient spoločne s Fisherovou informačnou metrikou vyjadruje:

$$\nabla E(f(x)|\theta) = F_{\theta}^{-1} \nabla_{\theta} E(f(x)|\theta), \quad (48)$$

kde Fisherova informačná matica  $F_{\theta}$  je predpokladom Hessovej matice z  $-\ln p$  a vyjadruje nezávislý vzťah na zvolenej parametrizácii. Kombináciou predchádzajúcich rovníc dostaneme:

$$\nabla E(f(x)|\theta) = F_{\theta}^{-1} E(f(x) \nabla_{\theta} \ln p(x|\theta)) = E(f(x) F_{\theta}^{-1} \nabla_{\theta} \ln p(x|\theta)) \quad (49)$$

**4.4.3.4 Stacionarita** Je pomerne ľahko vidieť, že aktualizčné rovnice CMA-ES spĺňajú nejaké stacionárne podmienky v tom, že sú v podstate objektívne. V prípade neutrálneho výberu, kde  $x_{1:\lambda} \sim \mathcal{N}(m_k, \sigma_k^2 C_k)$  zistíme, že:

$$E(m_{k+1}|m_k) = m_k \quad (50)$$



a za určitých ďalších počiatočných podmienok:

$$E(\log \sigma_{k+1} | \sigma_k) = \log \sigma_k \quad (51)$$

a ďalšou malou úpravou pre prípad, kedy je funkcia indikátora nulová, dostaneme:

$$E(C_{k+1} | C_k) = C_k \quad (52)$$

**4.4.3.5 Invarianty** Invarianty znamenajú rovnomerný výkon na triede objektívnych funkcií. Sú brané ako výhoda, pretože umožňujú generalizáciu a predvídanie správania sa algoritmu, a preto posilňujú význam empirických výsledkov získaných na jednotlivých funkciách. Nasledujúce invarianty boli stanovené v rámci CMA-ES.

- Invariant zachovávajúci rad pomocou transformácií objektívnych funkčných hodnôt  $f$ , a preto pre každú  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  je správanie identické na  $f : x \rightarrow g(h(x))$  pre všetky striktne rastúce  $g : \mathbb{R} \rightarrow \mathbb{R}$ .
- Mierkový invariant - to znamená, že pre všetky  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  je správanie nezávislé pre  $\alpha > 0$  pre objektívnu funkciu  $f : x \rightarrow h(\alpha x)$  pre dané  $\sigma_0 \propto 1/\alpha$  a  $m_0 \propto 1/\alpha$ .
- Invariant zachovávajúci rotáciu prehladavaného priestoru, a preto pre každú  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  a  $z \in \mathbb{R}^n$  sa správanie na  $f : x \rightarrow h(Rx)$  je nezávislé na ortogonálnej matici  $R$ , pre dané  $m_0 = R^{-1}z$ .

Každý parameter optimalizačnej metódy by mal zachovávať tieto invarianty, ale väčšina metód aj tak nevykazuje všetky z vyššie popísaných invariantov.

**4.4.3.6 Konvergenčia** Koncepčné úvahy o algoritme, ako napríklad mierkový invariant algoritmu, analýza jednoduchších evolučných stratégií alebo empirické dôkazy ukazujú, že algoritmus nad veľkou triedou funkcií rýchlo konverguje ku globálnemu optimu, označenému ako  $x^*$ . Na niektorých funkciách prebieha konvergenčia nezávisle na počiatočných podmienkach s pravdepodobnosťou 1. Na niektorých funkciách je pravdepodobnosť menšia ako 1 a typicky závisí na počiatočnom  $m_0$  a  $\sigma_0$ . Najrýchlejšia možná miera konvergenzie  $k$  pre priame vyhľadávacie metódy môže byť popísaná ako:

$$\|m_k - x^*\| \approx \|m_0 - x^*\| e^{-ck} \quad (53)$$

pre nejaké  $c > 0$ . Pre lepší popis sa používa:

$$\frac{1}{k} \sum_{i=1}^k \log \frac{\|m_i - x^*\|}{\|m_1 - x^*\|} = \frac{1}{k} \log \frac{\|m_k - x^*\|}{\|m_0 - x^*\|} \rightarrow -c < 0, k \rightarrow \infty, \quad (54)$$

alebo:

$$E \log \frac{\|m_k - x^*\|}{\|m_{k-1} - x^*\|} \rightarrow -c < 0, k \rightarrow \infty. \quad (55)$$

Toto znamená, že priemerná vzdialenosť k optimu sa znižuje v každej iterácii o konštantú vzdialenosť, menovite  $\exp(-c)$ . Konvergenčná rýchlosť  $c$  je zhruba  $0.1\lambda/n$  pre dané  $\lambda$ , a nie je oveľa väčšie ako dimenzia  $n$ . Dokonca pri optimálnom  $\sigma$  a  $C$  konvergenčná rýchlosť nepresahuje  $0.25\lambda/n$ .

**4.4.3.7 Interpretácia ako transformácia koordinačného systému** Použitie neekvivalentnej kovariančnej matice pre viacrozmerné normálne rozdelenie v evolučnej stratégii je ekvivalentné transformácii koordinačného systému vektorov podľa vzorkovacieho predpisu

$$x_i \approx m_k + \sigma_k \times \mathcal{N}(0, C_k) \quad (56)$$

$$x_i \approx m_k + \sigma_k \times C_k^{-1/2} m_k + \sigma_k \times \mathcal{N}(0, I) \quad (57)$$

môžu byť ekvivalentne vyjadrené v priestore ako:

$$C_k^{-1/2} x_i \approx C_k^{-1/2} m_k + \sigma_k \times \mathcal{N}(0, I). \quad (58)$$

Kovariančná matica definuje bijektívne zobrazenie pre všetky vektory riešení v priestore pre identickú kovariančnú maticu. Pretože aktualizčné rovnice CMA-ES sú invarianty lineárneho koordinačného systému, CMA-ES môže byť chápaná ako adaptívna procedúra aplikovaná na jednoduchú evolučnú stratégiu s identickou kovariančnou maticou.

#### 4.4.4 Výkon v praxi

V porovnaní s ostatnými evolučnými stratégiami je z pohľadu užívateľa CMA-ES algoritmus kvázi bezparametrický. Používateľ si musí zvoliť počiatočný bod riešenia  $m_0 \in \mathbb{R}^n$  a počiatočnú veľkosť kroku  $\sigma_0 > 0$ . Optionálne môže užívateľ zmeniť počet kandidátnych riešení  $\lambda$ , aby zmenil charakteristické vyhľadávacie správanie. Tiež by mal prispôbiť terminačné podmienky, aby čo najlepšie zodpovedali problému [44].

CMA-ES bolo empiricky úspešné v stovkách prípadov a je považované za užitočné v prípadoch nekonvexných, neseparovateľných, podmienených multimodálnych objektívnych funkciách. Dimenzia prehľadávaného priestoru sa zvyčajne pohybuje od dva po niekoľko stoviek. V prípade “black box” problémov, kde gradienty nie sú dostupné a výpočty funkčných hodnôt sú jediné dostupné, CMA-ES metóda bola prekonaná inými metódami za nasledujúcich podmienok:

- na málo dimenzionálnych funkciách,  $n < 5$ , po použití Nelder-Meadovej metódy [38],
- na separovateľných funkciách s alebo bez závislosti medzi funkčnými hodnotami v prípade multimodality alebo pri veľkých dimenziách pomocou diferenciálnej evolúcie,

- na konvexne-kvadratických funkciách s malou alebo strednou podmienenosťou Hessevej matice pomocou BFGS [39] alebo NEWUOA metódami [40],
- na funkciách, ktoré môžu byť vyrátané pomocou malého počtu výpočtov funkčných hodnôt,  $10n$ , pomocou NEWUOA metódy.

Na separovateľných funkciách môže byť nevýhoda to, že CMA-ES metóda nedokáže nájsť všetky dostupné riešenia. Na neseparovateľných funkciách, ktoré su podmienené alebo môže byť vyriešené po viac ako  $100n$  výpočtoch funkčných hodnôt, CMA-ES metóda ukazuje obrovskú prevahu.

#### 4.4.5 Rozšírenia

(1+1)CMA-ES [41] metóda generuje iba jedno kandidátne riešenie v každej iterácii, ktoré sa stane novou strednou hodnotou ak je lepšie ako aktuálna stredná hodnota. Pre  $c_c = 1$  je CMA-ES metóda blízka Gaussovej adaptácii. Niektoré prirodzené evolučné stratégie sú podobné CMA-ES pri špecifickom nastavení parametrov. Tieto evolučné stratégie nevyužívajú evolučné cesty ale aktualizujú rozptyly a kovariancie v kovariančnej matici. CMA-ES metóda bola tiež rozšírená na multiobjektívnu optimalizáciu MO-CMA-ES [42]. Ďalším dôležitým rozšírením je pridanie negatívnej aktualizácie kovariančnej matice, a toto rozšírenie sa volá aktívne CMA [43].

## 5 Metodológia a implementácia

Metodológia výpočtu distribúcie smeru vetra pozostáva z nasledujúcich krokov:

### 5.1 Metodológia von Misesovej distribúcie

1. **Vizualizácia** dostupných veterných dát napríklad pomocou veternej ruže alebo pomocou 2D grafu, v ktorom použijeme histogram smeru vetra. Veterná ruža sa používa predovšetkým v meteorológii a zobrazuje ako sú distribuované rýchlosti a smer vetra v určitom mieste. Distribúcia vetra je tu rozdelená do lúčov, pričom najdlhší lúč zobrazuje smer vetra s najväčšou frekvenciou. Histogram používa rozdelenie celkového počtu dát do rovnako širokých intervalov, a potom spočíta koľko prípadov spadá do jednotlivých kategórií.
2. **Odhad** počtu prevládajúcich smerov vetra  $N$ . Toto je dôležité, pretože potrebujeme vedieť koľko von Misesových distribúcií je nutné pomiešať, aby sme dokázali pokryť všetky prevládajúce smery vetra. Toto môžeme vyčítať z veternej ruže alebo histogramu.
3. **Rozdelenie** celého spektra smeru vetra:
  - Separácia dát do frekvenčných tried. Môžeme zvoliť rozdelenie napríklad po 10 stupňoch.
  - Rozdelenie frekvenčných tried do  $N$  sektorov rovnakej dĺžky.
4. **Odhad počiatočných parametrov**  $(\mu_j, \kappa_j, \omega_j)$  jednoduchej von Misesovej funkcie hustoty pravdepodobnosti pre každý z  $N$  sektorov:
  - *Prevládajúci smer vetra*  
Stredná hodnota smeru vetra  $\mu_j$  pre každý sektor je odhadovaná prostredníctvom

$$\mu_j = \arctan\left(\frac{s_j}{c_j}\right); \quad s_j \geq 0, c_j > 0, \quad (59)$$

$$\mu_j = \frac{\Pi}{2}; \quad s_j > 0, c_j = 0, \quad (60)$$

$$\mu_j = \Pi + \arctan\left(\frac{s_j}{c_j}\right); \quad c_j < 0, \quad (61)$$

$$\mu_j = \Pi; \quad s_j = 0, c_j = -1, \quad (62)$$

$$\mu_j = 2\Pi + \arctan\left(\frac{s_j}{c_j}\right); \quad s_j < 0, c_j > 0, \quad (63)$$

$$\mu_j = \frac{3\Pi}{2}; \quad s_j < 0, c_j = 0, \quad (64)$$

kde  $s_j$  a  $c_j$  reprezentujú priemerný sínus a kosínus pre všetky dáta v sektore nasledovne:

$$s_j = \frac{\sum_{i=1}^t n_i \sin(\theta_i)}{\sum_{i=1}^t n_i} \quad (65)$$

$$c_j = \frac{\sum_{i=1}^t n_i \cos(\theta_i)}{\sum_{i=1}^t n_i}, \quad (66)$$

kde  $t$  je počet frekvenčných tried v jednom sektore a  $n_i$  je počet pozorovaní v jednotlivých triedach.

- *Koncentračný parameter*

Pre každý sektor sa koncentračný parameter  $\kappa_j$  vypočítava iteratívne podľa nasledujúcej rovnice:

$$\frac{I_1(\kappa_j)}{I_0(\kappa_j)} = \left[ s_j^2 + c_j^2 \right]^{1/2}, \quad (67)$$

kde  $I_1(\kappa_j)$  je modifikovaná Besselova funkcia prvého radu. Analyticky je definovaná ako:

$$I_1(\kappa_j) = \frac{1}{\sqrt{2\Pi}} \int_0^{2\Pi} \cos \theta e^{\kappa_j \cos \theta} d\theta = \sum_{k=0}^{\infty} \frac{1}{\Gamma(k+2)\Gamma(k+1)} \left( \frac{\kappa_j}{2} \right)^{2k+1}, \quad (68)$$

kde  $\Gamma(n) = (n-1)!$  je klasická gamma funkcia. Koncentrácia  $\kappa_j$  môže byť tiež aproximovaná ako

$$\|\kappa_j\| = \{23.29041409 - 16.8617370(s_j^2 + c_j^2)^{0.25} - 17.4749884e^{-(s_j^2 + c_j^2)}\}^{-1}. \quad (69)$$

- *Váha*

Počiatkové váhy sú vyrátané podľa predpisu

$$\omega_j = \frac{\sum_{i=1}^t n_i}{\sum_{i=1}^T n_i}, \quad (70)$$

kde  $t$  je počet tried v jednom sektore,  $T$  je celkový počet frekvenčných tried, a  $n_i$  je počet pozorovaní v jednotlivých triedach.

## 5.2 Metodológia algoritmu CMA-ES

Predtým ako popíšeme implementáciu algoritmu je potrebné popísať pár základných operácií.

### 5.2.1 Dekompozícia pozitívne definitnej matice na vlastné vektory

Symetická, pozitívne definitná matica  $C \in \mathbb{R}^{n \times n}$  je charakterizovaná tak, že pre všetky  $x \in \mathbb{R}^n \setminus \{0\}$  platí  $x^T C x > 0$  [44]. Matica  $C$  má ortonormálnu bázu vlastných vektorov  $B = [b_1, \dots, b_n]$  s korešpondujúcimi vlastnými číslami  $d_1^2, \dots, d_n^2 > 0$ .

Pre dekompozíciu  $C$  platí:

$$C = B D^2 B^T, \quad (71)$$

kde

- $B$  je ortogonálna matica,  $B^T B = B B^T = I$ . Stĺpce matice tvoria ortonormálnu bázu vlastných vektorov.
- $D^2 = D D = \text{diag}(d_1, \dots, d_n)^2 = \text{diag}(d_1^2, \dots, d_n^2)$  je diagonálna matica s vlastnými číslami  $C$  na diagonále.
- $D = \text{diag}(d_1, \dots, d_n)$  je diagonálna matica s odmocninami vlastných čísel  $C$  na diagonále.

Dekompozícia matice je unikátna, okrem toho, že stĺpce matice  $\mathbf{B}$  a permutácie týchto stĺpcov môžu dávať rozdielne vlastné čísla.

V praxi musí byť výpočet matíc  $\mathbf{B}$  a  $\mathbf{D}$  prevedený v maximálnom počte  $\max(1, \lfloor 1/(10n * (c_1 + c_\mu)) \rfloor)$  generácií.

### 5.2.2 Viacrozmerné normálne rozdelenie

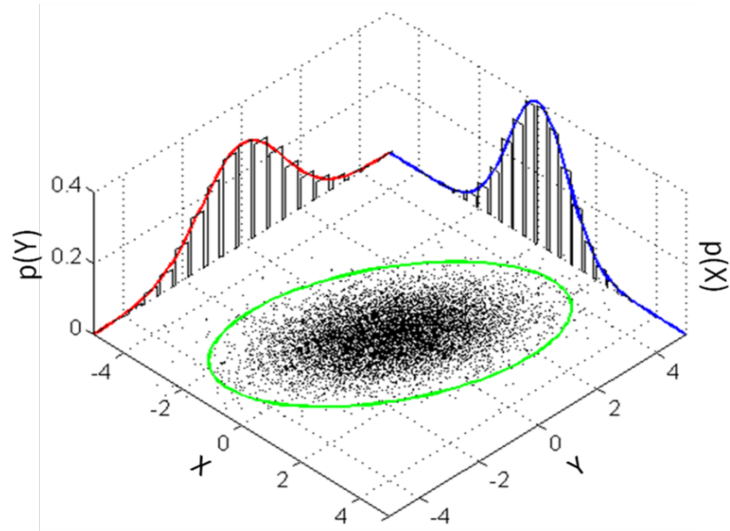
Viacrozmerné normálne rozdelenie  $\mathcal{N}(m, C)$  má unimodálnu, zvonovú hustotu, kde vrch zvona korešponduje so strednou hodnotou  $m$  [44]. Distribúcia je unikátne určená strednou hodnotou  $m \in \mathbb{R}^n$  a symetrickou pozitívne definitnou kovariančnou maticou  $C \in \mathbb{R}^{n \times n}$ . Normálna distribúcia  $\mathcal{N}(m, C)$  môže byť zapísaná pomocou

$$m \sim \mathcal{N}(0, C) \quad (72)$$

$$m \sim C^{\frac{1}{2}} \mathcal{N}(0, I) \quad (73)$$

$$m \sim B D B^T \mathcal{N}(0, I) \quad (74)$$

$$m \sim B D \mathcal{N}(0, I). \quad (75)$$



Obr. 6: Funkcia hustoty pravdepodobnosti viacrozmerného normálneho rozdelenia [51].

### 5.2.3 Náhodná black box optimalizácia

Použijeme scenár, kde chceme minimalizovať objektívnu funkciu:

$$f : \mathbb{R}^n \rightarrow \mathbb{R} \quad (76)$$

$$f : x \rightarrow f(x). \quad (77)$$

Cielom je nájsť jeden alebo viacero bodov  $x \in \mathbb{R}^n$  s funkčnou hodnotou  $f(x)$  tak malou, ako je to len možné. Black box optimalizácia opisuje situáciu, kedy jedinou dostupnou informáciou sú iba funkčné hodnoty hľadaných bodov. Hľadané body môžu byť vybraté náhodne. Definujeme *cenu vyhľadávania* ako počet vykonaných výpočtov funkčných hodnôt. Táto výkonnostná miera zachováva cenu vyhľadávania spoločne so zachovaním funkčných hodnôt.

### 5.2.4 Hessove a kovariančné matice

Uvažujme o konvexne-kvadratickej objektívnej funkcii  $f_H : x \rightarrow \frac{1}{2}x^T H x$ , kde Hessova matica  $H \in \mathbb{R}^{n \times n}$  je pozitívne definitná matica. Pre danú distribúciu  $\mathcal{N}(m, C)$  existuje blízka relácia medzi  $H$  a  $C$ : Ak nastavíme  $C = H^{-1}$  funkciu  $f_H = \frac{1}{2}x^T x = \frac{1}{2} \sum_i x_i^2$  môžeme zameniť  $H = I$  za  $C = I$ . To znamená, že pre konvexne-kvadratické funkcie môžeme ekvivalentne používať kovariančné alebo Hessove matice.



### 5.3 Implementácia algoritmu CMA-ES

V tejto podkapitole sumarizujeme algoritmus  $(\mu/\mu_W, \lambda)$ -CMA-ES a konkrétny popis implementácie, ktorá bola použitá v tejto práci [44].

Prvý krokom implementácie je nastavenie stavových premenných algoritmu:

- $\lambda$  - veľkosť populácie, platí  $\lambda \geq 2$  a jej počiatková hodnota je:

$$\lambda = 4 + \lfloor 3 \ln n \rfloor \quad (78)$$

- $\mu$  - počet rodičov, platí  $\mu \leq \lambda$  a jeho počiatková hodnota je:

$$\mu = \lfloor \lambda/2 \rfloor \quad (79)$$

- $w_{i=1\dots\mu}$  - pozitívne váhové koeficienty pre rekombináciu, platí  $\sum_{i=1}^{\mu} w_i = 1$  a počiatkové hodnoty sú:

$$w_i = \frac{\dot{w}_i}{\sum_{j=1}^{\mu} \dot{w}_j^+}, \dot{w}_i = \ln(\lambda/2 + 1/2) - \ln i \quad (80)$$

- $c_{\sigma}$  - rýchlosť učenia pre kumuláciu pre kontrolu veľkosti kroku, platí  $c_{\sigma} < 1$  a počiatková hodnota je:

$$c_{\sigma} = \frac{\mu_{eff} + 2}{n + \mu_{eff} + 5} \quad (81)$$

kde  $\mu_{eff} = \frac{1}{\sum_{i=1}^{\mu} w_i^2}$

- $d_{\sigma}$  - útlmový parameter pre aktualizáciu veľkosti kroku,  $d_{\sigma} \approx 1$  a počiatková hodnota je:

$$d_{\sigma} = 1 + 2 * \max \left( 0, \sqrt{\frac{\mu_{eff} - 1}{n + 1}} - 1 \right) + c_{\sigma} \quad (82)$$

- $c_c$  - rýchlosť učenia pre kumuláciu pre rank-one aktualizáciu kovariančnej matice, platí  $c_c \leq 1$  a počiatková hodnota je:

$$c_c = \frac{4 + \mu_{eff}/n}{n + 4 + 2 * \mu_{eff}/n} \quad (83)$$

- $c_1$  - rýchlosť učenia pre rank-one aktualizáciu kovariančnej matice,  $c_1 \leq 1 - c_{\mu}$  a počiatková hodnota je:

$$c_1 = \frac{2}{(n + 1.3)^2 + \mu_{eff}} \quad (84)$$

- $c_{\mu}$  - rýchlosť učenia pre rank- $\mu$  aktualizáciu kovariančnej matice,  $c_{\mu} \leq 1 - c_1$ , počiatková hodnota je:

$$c_{\mu} = \min \left( 1 - c_1, \alpha_{\mu} * \frac{\mu_{eff} - 2 + 1/\mu_{eff}}{(n + 2)^2 + \alpha_{\mu} * \mu_{eff}/2} \right), \alpha_{\mu} = 2 \quad (85)$$

Ďalej inicializujeme parametre nasledovne: evolučné cesty  $p_\sigma = 0$  a  $p_c = 0$ , kovariančnú maticu  $C = I$  a počet generácií  $g = 0$ .

Teraz nasleduje iteračný priebeh algoritmu, kde v každej iterácii cyklu prebehnú nasledujúce operácie:

- Vzorkovanie novej populácie bodov pre  $k = 1, \dots, \lambda$

$$\mathbf{z}_k \sim \mathcal{N}(0, I) \quad (86)$$

$$\mathbf{y}_k = \mathbf{B}\mathbf{D}\mathbf{z}_k \sim \mathcal{N}(0, C) \quad (87)$$

$$\mathbf{x}_k = \mathbf{m} + \sigma\mathbf{y}_k \sim \mathcal{N}(\mathbf{m}, \sigma^2 C). \quad (88)$$

Inými slovami, najskôr sa vytvorí vektor  $\mathbf{z}_k$  z viacrozmerného normálneho rozdelenia s nulovou strednou hodnotou a jednotkovou kovariančnou maticou. Vektor  $\mathbf{y}_k$  vznikne prenasobením vektora  $\mathbf{z}_k$  vlastnými vektormi (matica  $\mathbf{B}$ ) a vlastnými číslami (matica  $\mathbf{D}$ ) matice  $\mathbf{C}$ . Populácia, ktorá je reprezentovaná vektorom  $\mathbf{x}_k$ , vznikne pričítaním strednej hodnoty distribúcie k vektoru  $\mathbf{y}_k$ , ktorý je prenasobený rozptylom.

- Selekcia a rekombinácia

$$\langle \mathbf{y}_w \rangle = \sum_{i=1}^{\mu} w_i \mathbf{y}_{i:\lambda} \quad (89)$$

$$\mathbf{m} \leftarrow \mathbf{m} + \sigma \langle \mathbf{y}_w \rangle = \sum_{i=1}^{\mu} w_i \mathbf{x}_{i:\lambda} \quad (90)$$

Novú strednú hodnotu populácie dostaneme tak, že pôvodný vektor populácie vynásobíme váhovým vektorom.

- Kontrola veľkosti kroku

$$\mathbf{p}_\sigma \leftarrow (1 - c_\sigma) * \mathbf{p}_\sigma + \sqrt{c_\sigma(2 - c_\sigma) * \mu_{eff} * \mathbf{C}^{-\frac{1}{2}}} * \langle \mathbf{y}_w \rangle \quad (91)$$

$$\sigma \leftarrow \sigma \times \exp \left( \frac{c_\sigma}{d_\sigma} \left( \frac{\|\mathbf{p}_\sigma\|}{E \|\mathcal{N}(0, I)\|} - 1 \right) \right) \quad (92)$$

Evolučnú cestu  $p_\sigma$  sa aktualizuje pomocou aktuálnej evolučnej cesty a spätného časového horizontu  $c_\sigma$ , ktorý sa pohybuje v rozmedzí  $\sqrt{N}$  a  $N$ .  $\sqrt{c_\sigma(2 - c_\sigma) * \mu_{eff}}$  je normalizačná konštanta.

Pri aktualizácii  $\sigma$  porovnáваме či predpokladná dĺžka vyhovuje  $E \|\mathcal{N}(0, I)\|$ , teda Euklidovskej norme vektora z normálneho rozdelenia, ktorá je navyše utlmená útlmovým parametrom  $d_\sigma$ .

- Adaptácia kovariančnej matice

$$\mathbf{p}_c \leftarrow (1 - c_c)\mathbf{p}_c + h_\sigma \sqrt{c_c(2 - c_c) * \mu_{eff}} \langle \mathbf{y}_w \rangle \quad (93)$$

$$\mathbf{C} \leftarrow (1 - c_1 - c_\mu)\mathbf{C} + c_1(\mathbf{p}_c\mathbf{p}_c^T + \delta(h_\sigma)\mathbf{C}) + c_\mu \sum_{i=1}^{\mu} w_i \mathbf{y}_{i:\lambda} \mathbf{y}_{i:\lambda}^T \quad (94)$$

Evolučná cesta  $p_c$  sa aktualizuje pomocou aktuálnej evolučnej cesty a spätného časového horizontu  $c_\sigma$ , ktorý pokrýva zhruba 63% celkovej váhy.  $\sqrt{c_\sigma(2 - c_\sigma) * \mu_{eff}}$  je normalizačná konštanta.  $h_\sigma$  utlmuje túto evolučnú cestu aby sa zabránilo príliš rýchlej konvergencii. Nadobúda hodnoty 1 ak platí

$$\frac{\|p_\sigma\|}{\sqrt{1 - (1 - c_\sigma)^{2(g+1)}}} < (1.4 + \frac{2}{n+1})E \|\mathcal{N}(0, I)\|, \quad (95)$$

kde  $g$  je číslo generácie. Ak predpis 95 neplatí,  $h_\sigma$  nadobúda hodnotu 0.  $\delta(h_\sigma)$  nadobúda hodnotu  $(1 - h_\sigma)c_c(2 - c_c)$ .

### 5.3.1 Terminačné kritériá

Pre zastavenie iteračného cyklu je nutné použiť terminačné kritéria. V našom prípade boli použité nasledovné:

- *NoEffectAxis* - zastavenie ak pridanie vektora 0.1 - štandardnej odchýlky k hocijakej hlavnej osi matice  $\mathbf{C}$  nezmení strednú hodnotu  $\mathbf{m}$ .
- *NoEffectCoord* - zastavenie ak pridanie vektora 0.2 - štandardnej odchýlky k hocijakému jednému koordinátu nezmení strednú hodnotu  $\mathbf{m}$ .
- *EqualFunValues* - zastavenie ak je rozsah najlepších funkčných hodnôt objektívnej funkcie pre posledných  $10 + \lceil 30n/\lambda \rceil$  generácií nulový.
- *Stagnation* - ukladáme si históriu najlepšej a priemernej fitness hodnoty v každom iteračnom kroku za najviac 20% ale najmenej  $120 + 30n/\lambda$  a nie viac ako 20 000 iterácií. Zastavujeme ak v oboch históriách priemer posledných 30% funkčných hodnôt nie je lepší ako priemer prvých 30%.
- *ConditionCov* - zastavenie ako podmienené číslo kovariančnej matice presiahne  $10^{14}$ .
- *To1XUp* - zastavenie ak  $\sigma \times \max(diag(\mathbf{D}))$  sa zvýšil o viac ako  $10^4$ .
- *To1Fun* - zastavenie ak rozsah najlepších funkčných hodnôt objektívnej funkcie v posledných  $10 + \lceil 30n/\lambda \rceil$  generáciách a všetkých funkčných hodnôt v posledných generáciách pod *To1Fun*. *To1Fun* sa nastavuje na  $10^{-12}$ .

- *To1X* - zastavenie ak štandardná odchýlka normálneho rozdelenia je menšia vo všetkých koordinátoch a  $\sigma \mathbf{p}_c$  je menšie ako *To1X* vo všetkých komponentoch. *To1X* sa nastavuje na  $10^{-12} * \sigma$ .

## 5.4 Implementácia programu

Pre implementáciu programu bol zvolený programovací jazyk C++. Aplikácia mala byť podľa zadania nezávislá, a preto neobsahuje žiadne knižnice tretích strán. Aplikácia je implementovaná nasledovne:

Pred spustením samotného evolučného cyklu najskôr dôjde k inicializácii stavových premenných algoritmu CMA-ES. Dôjde teda k inicializácii premenných  $\lambda$ ,  $\mu$ ,  $w_{i=1..\mu}$ ,  $c_\sigma$ ,  $\mu_{eff}$ ,  $d_\sigma$ ,  $c_c$ ,  $c_1$ ,  $c_\mu$  na také počiatočné hodnoty, ako je to popísané v predchádzajúcej sekcii. V prípade váh je implementovaná podpora 3 rôznych typov váh:

- **LINEAR\_WEIGHTS** -  $\sum_i^\mu w_i = \mu - i$ ,
- **EQUAL\_WEIGHTS** -  $\sum_i^\mu w_i = 1$ ,
- **LOG\_WEIGHTS** -  $\sum_i^\mu w_i = \log(mu + 1) - \log(i + 1)$ .

Pre testovanie aplikácie však boli použité iba **LOG\_WEIGHTS**. Potom začína samotný iteračný priebeh algoritmu, kde sa vždy na začiatku cyklu skontrolujú terminačné kritéria, ktoré sú popísané vyššie. Ďalej nasleduje ovzorkovanie populácie pomocou viacrozmerného normálneho rozdelenia, čo je tiež popísané vyššie. Tiež tu dochádza k testovaniu strednej odchýlky populácie, aby neprekročila nastavené limity a k prevzorkovaniu jedincov, ktorý tieto limity prekročili. Potom dochádza k ohodnoteniu jedincov pomocou fitness funkcie. V nasledujúcom kroku dochádza k aktualizácii distribúcie. Je to kľúčová procedúra, v ktorej najskôr dôjde k usporiadaniu jedincov podľa indexov, a tiež uloženiu týchto fitness hodnôt. História fitness hodnôt sa používa pri terminácii algoritmu. Nasleduje výpočet strednej hodnoty populácie  $m$ , výpočet evolučnej cesty  $p_\sigma$  a evolučnej cesty  $p_c$ . Na základe týchto dvoch hodnôt potom dochádza k aktualizácii kovariančnej matice  $C$  a nakoniec k aktualizácii  $\sigma$ .

Po skončení iteračného cyklu program zobrazí na výstup najlepšiu fitness hodnotu, ktorá bola dosiahnutá, a tiež najlepšiu dosiahnutú strednú hodnotu populácie. Pri testovaní algoritmu je použitá kruhová distribúcia, ktorej interval  $[0, 2\pi]$  je rozdelený na 36 rovnakých sektorov po 10 stupňoch. Preto na výstupe je najskôr zobrazených 36 analyticky odhadnutých hodnôt pre jednotlivé sektory a potom ďalších 36 hodnôt optimalizovaných pomocou evolučného algoritmu. Ďalej sú vypísané výsledky štatistických testov, konkrétne: Pearsonov  $\chi^2$  test, RMSE a  $R^2$ .

V poslednej časti sú vypísané hodnoty von Misesovej distribúcie:  $\theta$  (stredná hodnota),  $\kappa$  (rozptyl),  $\omega$  (váha) pre všetkých 6 frekvenčných tried. Nakoniec sú vypísané jednotlivé hodnoty z in-

tervalu  $[0, 2\pi]$  s odstupom 0,001, teda od hodnoty 0,000 po hodnotu 6,283. Pre každú túto hodnotu je vypísaná analyticky odhadnutá hodnota a hodnota optimalizovaná pomocou evolučného algoritmu. Tieto hodnoty sú potom použité v histogramoch, ktoré sú zobrazené v kapitole 6.

#### 5.4.1 Organizácia kódu

Algoritmus popísaný v tejto práci je implementovaný v rámci väčšej aplikácie, ktorá obsahuje ďalšie biologicky inšpirované algoritmy pre optimalizáciu ročného rozdelenia vetra. V tomto prehľade sú preto uvedené súbory, ktoré boli implementované v rámci tejto práce.

- *cCMAES.h* - hlavičkový súbor.
- *cCMAESParameters.h* - súbor obsahuje metódy pre inicializáciu stavových parametrov algoritmu CMA-ES.
- *cCMAESRandom.h* - súbor obsahuje pseudo náhodný generátor čísel.
- *cCMAES.cpp* - súbor obsahuje rozhranie pre prácu s algoritmom CMA-ES. Nachádzajú sa tu kľúčové procedúry, napríklad, pre aktualizáciu distribúcie a adaptácia kovariančnej matice.
- *main.cpp* - v súbore boli prevedené úpravy, aby sa dokázal spustiť algoritmus CMA-ES.

## 5.5 Štatistické testy

V práci sú použité 3 typy testov, ktoré rozhodujú o tom, či a ako veľmi sa pozorovaná frekvenčná distribúcia líši od teoretickej distribúcie. Určitý typ testu môže byť nastavený pri spúšťaní aplikácie.

### 5.5.1 Pearsonov chí-kvadrát test

Tento štatistický test sa dá použiť na súbory kategorických dát, ako napríklad histogramy. Výpočet nám udáva, ako je pravdepodobné, že pozorované rozdiely medzi súbormi vznikli náhodou. Výsledok je vypočítaný pomocou referencie na  $\chi^2$  distribúciu. Jeho vlastnosti boli prvýkrát skúmané Karlom Pearsonom v roku 1900 [45].

Hodnota testovej štatistiky je:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = N \sum_{i=1}^n p_i \left( \frac{O_i/N - p_i}{p_i} \right)^2, \quad (96)$$

kde:

- $\chi^2$  = Pearsonova kumulatívna testová štatistika, ktorá asymptoticky aproximuje  $\chi^2$  distribúciu,
- $O_i$  = počet pozorovaní typu  $i$ ,
- $N$  = celkový počet pozorovaní,
- $E_i = N p_i$  = predpokladaná frekvencia typu  $i$ , znamená, že pravdepodobnosť typu  $i$  v populácii  $N$  je  $p_i$ ,
- $n$  = počet frekvenčných tried.

### 5.5.2 Koeficient determinácie

Koeficient determinácie, označovaný ako  $R^2$  alebo  $r^2$ , je číslo, ktoré indikuje ako dobre dáta zodpovedajú štatistickému modelu. Ak sa  $R^2$  rovná 1, znamená to, že regresná priamka perfektne zodpovedá dátam. Ak sa  $R^2$  rovná 0 indikuje to, že priamka vôbec nezodpovedá dátam.

Dataset má  $n$  hodnôt označených ako  $y_1, \dots, y_n$  a každej je priradená predikovaná hodnota  $f_1, \dots, f_n$ . Definujeme reziduá ako  $e_i = y_i - f_i$ . Ak  $\bar{y}$  je stredná hodnota pozorovaných dát [46]:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad (97)$$

potom môže byť variabilita datasetu meraná pomocou troch súm druhých mocnín [46]:

- Celková suma druhých mocnín

$$SS_{tot} = \sum_i (y_i - \bar{y})^2, \quad (98)$$

- Regresná suma druhých mocnín

$$SS_{reg} = \sum_i (f_i - \bar{y})^2, \quad (99)$$

- Suma druhých mocnín reziduí

$$SS_{res} = \sum_i (y_i - f_i)^2 = \sum_i e_i^2. \quad (100)$$

Najvšeobecnejšia definícia koeficientu determinácie je potom:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}. \quad (101)$$

### 5.5.3 Root-mean-square odchýlka

Odchýlka odmocniny priemeru alebo chyba odmocniny priemeru je často používaná miera pre meranie rozdielov medzi hodnotami predpovedanými modelom a hodnotami pozorovanými v skutočnosti. RSME predstavuje štandardnú odchýlku vzorky rozdielov medzi predikovanými a pozorovanými hodnotami. Individuálne rozdiely sú nazývané reziduá ak sú výpočty prevádzané priamo nad vzorkou dát, v opačnom prípade sú nazývané chyby predikcie.

RSME vzhľadom na odhadovaný parameter  $\theta$  je odhadovaná ako odmocnina mocniny chyby strednej hodnoty [47]:

$$RMSE(\theta) = \sqrt{MSE(\theta)} = \sqrt{E((\hat{\theta} - \theta)^2)}. \quad (102)$$

RMSE predikovaných hodnôt  $\hat{y}_t$  regresne závislej premennej  $y$  je vypočítaná pre  $n$  rozdielných predikcií ako odmocnina strednej hodnoty mocnín odchýliek [47]:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y)^2}{n}}. \quad (103)$$

## 6 Testovanie

Testovanie algoritmu prebiehalo v dvoch fázach. V prvej fáze bol algoritmus testovaný na známych matematických optimalizačných úlohách a v druhej fáze bol otestovaný na reálnych smerových dátach vetra. V tejto kapitole sú opísane obidve testovania spoločne s ich výsledkami.

### 6.1 Testovanie CMA-ES algoritmu

Algoritmus bol pred nasadením na reálne dáta otestovaný na matematických funkciách, aby sa ukázalo, či je správne implementovaný a predovšetkým optimalizovaný na daný typ úlohy. Bolo vybraných niekoľko matematických funkcií, ktoré budú v nasledujúcej sekcii popísané. Funkcie su prebraté z [48].

Objektívna funkcia č. 1: **prvá de Jongova funkcia**. Matematický predpis:

$$\sum_{i=1}^D x_i^2 \quad (104)$$

Objektívna funkcia č. 2: **druhá de Jongova funkcia** inak nazývaná tiež **Rosenbrockovo sedlo**. Matematický predpis:

$$\sum_{i=1}^{D-1} 100(x_i^2 - x_{i+1})^2 + (1 - x_i)^2 \quad (105)$$

Objektívna funkcia č. 3: **tretia de Jongova funkcia**. Matematický predpis:

$$\sum_{i=1}^D |x_i| \quad (106)$$

Objektívna funkcia č. 4: **štvrtá de Jongova funkcia**. Matematický predpis:

$$\sum_{i=1}^D i * x_i^4 \quad (107)$$

Objektívna funkcia č. 5: **Rastriginova funkcia**. Matematický predpis:

$$2D \sum_{i=1}^D x_i^2 - 10 \cos(2\pi x_i) \quad (108)$$

Objektívna funkcia č. 6: **Schwefelova funkcia**. Matematický predpis:

$$\sum_{i=1}^D -x_i \sin(\sqrt{|x_i|}) \quad (109)$$



Objektívna funkcia č. 7: **Griewangkova funkcia**. Matematický predpis:

$$1 + \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) \quad (110)$$

Objektívna funkcia č. 8: **Sínusová obáľková sinusoidálna funkcia**. Matematický predpis:

$$- \sum_{i=1}^{D-1} \left( 0.5 + \frac{\sin(x_i^2 + x_{i+1}^2 - 0.5)^2}{(1 + 0.001(x_i^2 + x_{i+1}^2))^2} \right) \quad (111)$$

Objektívna funkcia č. 9: **Rozťahnutá sinusoidálna V funkcia**. Matematický predpis:

$$\sum_{i=1}^{D-1} \left( \sqrt[4]{(x_i^2 + x_{i+1}^2)} \sin(50 \sqrt[10]{(x_i^2 + x_{i+1}^2)^2} + 1) \right) \quad (112)$$

Objektívna funkcia č. 10: **Ackleyho funkcia I**. Matematický predpis:

$$\sum_{i=1}^{D-1} \left( \frac{1}{e^5} \sqrt{(x_i^2 + x_{i+1}^2)} + 3(\cos(2x_i) + \sin(2x_{i+1})) \right) \quad (113)$$

Objektívna funkcia č. 11: **Ackleyho funkcia II**. Matematický predpis:

$$\sum_{i=1}^{D-1} \left( 20 + e - \frac{20}{e^{0.2 \sqrt{\frac{(x_i^2 + x_{i+1}^2)}{2}}}} - e^{0.5(\cos(2\pi x_i) + \cos(2\pi x_{i+1}))} \right) \quad (114)$$

Objektívna funkcia č. 12: **Pláto vajec**. Matematický predpis:

$$\sum_{i=1}^{D-1} \left( -x_i \sin(\sqrt{|x_i - x_{i+1} - 47|}) - (x_{i+1} + 47) \sin(\sqrt{|x_{i+1} + 47 + \frac{x_i}{2}|}) \right) \quad (115)$$

Objektívna funkcia č. 13: **Ranova funkcia**. Matematický predpis:

$$\begin{aligned} \sum_{i=1}^{D-1} (x_i \sin(\sqrt{|x_{i+1} + 1 - x_i|}) \cos(\sqrt{|x_{i+1} + 1 + x_i|}) \\ + (x_{i+1} + 1) \cos(\sqrt{|x_{i+1} + 1 - x_i|}) \sin(\sqrt{\sqrt{|x_{i+1} + 1 + x_i|}})) \end{aligned} \quad (116)$$

Objektívna funkcia č. 14: **Patologická funkcia**. Matematický predpis:

$$\sum_{i=1}^{D-1} \left( 0.5 + \frac{\sin(\sqrt{100x_i^2 - x_{i+1}^2}^2 - 0.5)}{(1 + 0.001(x_i^2 - 2x_i x_{i+1} + x_{i+1}^2)^2)} \right) \quad (117)$$

Objektívna funkcia č. 15: **Michalewiczova funkcia**. Matematický predpis:

$$\sum_{i=1}^{D-1} \left( -1 \left( \sin(x_i) \sin\left(\frac{x_i^2}{\Pi}\right)^{20} + \sin(x_{i+1}) \sin\left(\frac{2x_i^2}{\Pi}\right)^{20} \right) \right) \quad (118)$$

Objektívna funkcia č. 16: **Mastersova funkcia**. Matematický predpis:

$$\sum_{i=1}^{D-1} \left( e^{\frac{-(x_i^2 + x_{i+1}^2 + 0.5x_i x_{i+1})}{8}} \cos(4\sqrt{x_i^2 + x_{i+1}^2 + 0.5x_i x_{i+1}}) \right) \quad (119)$$

Z praktického hľadiska nie je potrebné uvádzať výsledky testov na uvedených objektívnych funkciách, pretože to nebolo cieľom práce. Funkcie boli testované iba preto, aby sa ukázalo či je algoritmus dobre implementovaný a optimalizovaný pre daný typ úlohy. Testy ukázali, že algoritmus je dobre optimalizovaný a vracia na testovacie funkcie očakávané výsledky. Tým sa potvrdilo, že implementácia algoritmu je dobrá a mohol byť otestovaný na reálnych smerových dátach, čo je popísané v nasledujúcej časti.

## 6.2 Testovanie výslednej aplikácie

K zlepšeniu výkonu CMA-ES algoritmu môžu byť odhadnuté počiatočné parametre zmesi von Misesovej distribúcie tak, že odhadneme počiatočné parametre jednoduchých von Misesových distribúcií pre jednotlivé rovnako vzdialené dátové sektory. Najskôr sa musí odhadnúť počet prevládajúcich smerov vetra. Tento parameter môžeme vyčítať z veternej ruže alebo z histogramu frekvencie dát. Po vzoru ostatných prác zaoberajúcich sa naším problémom, bol počet prevládajúcich smerov stanovený na  $N = 6$ .

Všetky dáta boli spočítané a rozdelené do frekvenčných tried po 10 stupňov a vytvorili tak frekvenčný histogram. Triedy boli potom rozdelené do 6 počiatočných rovnako vzdialených sektorov po 60 stupňov. Boli vypočítané parametre funkcie hustoty pravdepodobnosti jednoduchých von Misesových distribúcií pre každý sektor použitím už predtým popísanej metodiky.

Čo sa týka algoritmu CMA-ES neexistujú žiadne odporúčenia ako optimálne implementovať tento algoritmus pre náš problém. Preto bolo zvolené nasledujúce nastavenie: zmes  $k$  von Misesových distribúcií je reprezentovaná kandidátnym vektorom  $\mathbf{v} = (v_1, \dots, v_n)$ ,  $v_i \in [-1, 1]$ , ktorý je zložený z 3 vektorov  $\mathbf{v} = (\mu, \kappa, \omega)$ . Vektor  $\mu$  zahŕňa škálovanie  $v_i$ ,  $i \in 1, \dots, k$  do hodnôt  $[0, 2\Pi]$  a vektor  $\kappa$  vyžaduje škálovanie  $v_i$ ,  $i \in k + 1, \dots, 2k$  do hodnôt  $[0, 700]$ . Horná hranica intervalu  $\kappa$  bola vybratá s ohľadom na precíznosť použitého numerického algoritmu. Obe škálovania sa aplikujú tak, aby spĺňali požiadavky vektora  $\omega$ , ktorý je definovaný takto:

$$v_{ij} = 1, i = 2k + 1 \quad (120)$$

$$v_i \left( 1 - \sum_{j=1}^{i-1} \omega_j \right) > 1, i = 2k + 1 + j. \quad (121)$$

To garantuje, že suma všetkých váh je rovná 1 a nevyžaduje žiadne ďalšie obmedzenia pri práci s vektorom kandidátnych riešení. Všetky vektory vytvorené pri optimalizačnom procese sú validnými kandidátnymi riešeniami a reprezentujú zmes von Misesových distribúcií. Počiatočné hodnoty všetkých 3 vektorov  $\mu, \kappa, \omega$  sú analyticky odhadnuté pomocou metodiky popísanej v kapitole 5. Fitness hodnota kandidátnych riešení je ohodnocovaná pomocou rôznych testov a je to vždy uvedené pri každom testovaní.

Parametre CMA-ES algoritmu boli zvolené nasledovne:

- **Dimenzia**  $N = 100$
- **Veľkosť populácie**  $\lambda = 17$
- **Počet rodičov**  $\mu = 8$
- **Počet iterácií**  $= 561653$

### 6.2.1 Drahany

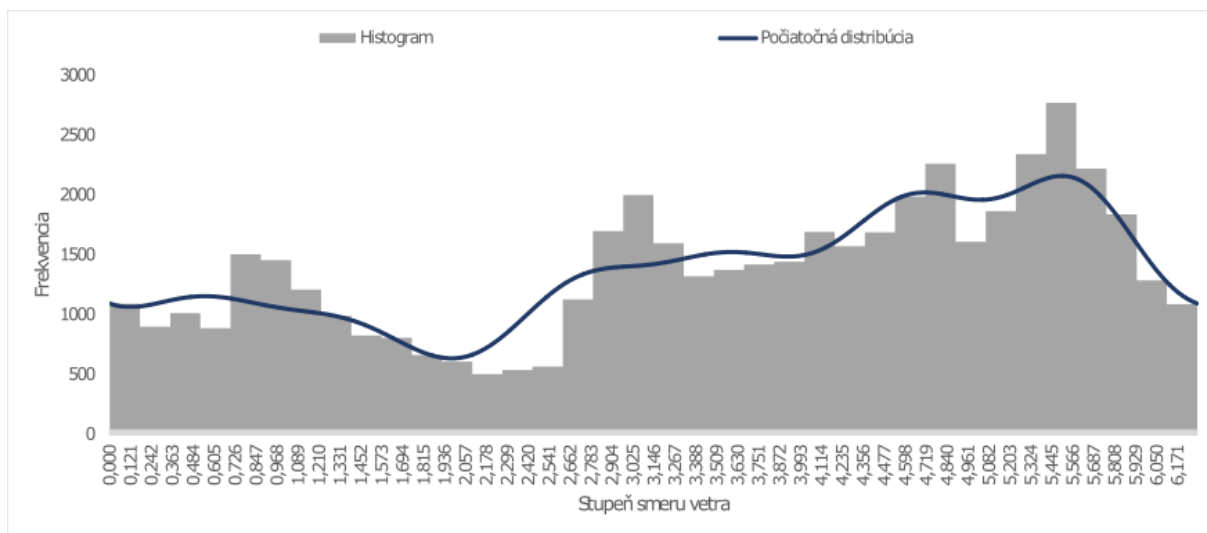
Dáta pochádzajú z veternej elektrárne Drahany, ktoré je vybudovaná neďaleko rovnako nazvanej obce. Obec leží v Olomouckom kraji v Českej republike. Elektráreň bola daná do prevádzky v roku 2006 a je v nej inštalovaná iba jedna veterná turbína. Všetky parametre počiatočnej zmesi von Misesových distribúcií sú popísané v tabuľke 1.

Tabuľka 1: Počiatočné parametre jednotlivých von Misesových distribúcií - Drahany

N	$\kappa_j$	$\mu_j$	$\omega_j$
1	4.62489	0.482081	0.137725
2	5.0164	1.41417	0.102798
3	5.86475	2.68914	0.129398
4	4.64526	3.58677	0.177715
5	5.31318	4.63724	0.220548
6	5.79601	5.58716	0.231816

Obrázok 7 zobrazuje normalizovaný dátový frekvenčný histogram spoločne s počiatočným modelom. Na týchto dátach bola prevedená séria experimentov, pri ktorej bol algoritmus nastavený tak ako je to popísané vyššie. Každý experiment bol opakovaný 30 krát vzhľadom na stochastické správanie sa algoritmu CMA-ES. Ako fitness funkcia bol použitý štatistický test RMSE, a to kvôli tomu, aby sa výsledky mohli porovnať s už existujúcimi výsledkami. Existujúce výsledky boli dodané vedúcim práce, kde v rámci jeho experimentov bol použitý na spracovanie rovnakých dát algoritmus DE a fitness funkcia RMSE. Preto bude možné porovnať výsledky algoritmus CMA-ES s algoritmom DE. Pri experimentoch sa algoritmus vždy zastavil po dosiahnutí maximálneho počtu iterácií. V tabuľke 2 je zobrazený najlepší, najhorší a priemerný výsledok RMSE testu a taktiež jeho porovnanie s algoritmom DE. Na obrázku 8 môžeme vi-

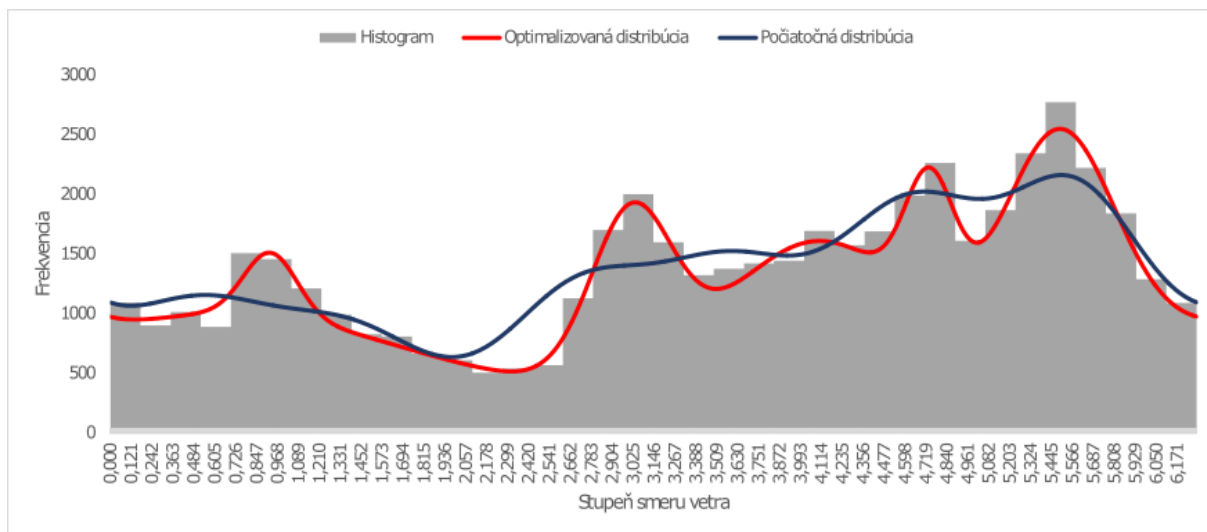
dieť optimalizovanú distribúciu pre smerové dáta vetra pomocou algoritmu CMA-ES. Obrázok znázorňuje situáciu pre najlepší prípad, ktorý bol dosiahnutý teda pre hodnotu  $RMSE = 87,51$ .



Obr. 7: Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Drahany

Tabuľka 2: Výsledky optimalizácie distribúcie smeru vetra pre Drahany

Algoritmus	najlepší	priemer	najhorší
(8,17)-CMA-ES	87,51	133,33	189,14
DE/1/rand	78,97	91,90	123,36



Obr. 8: Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Drahany

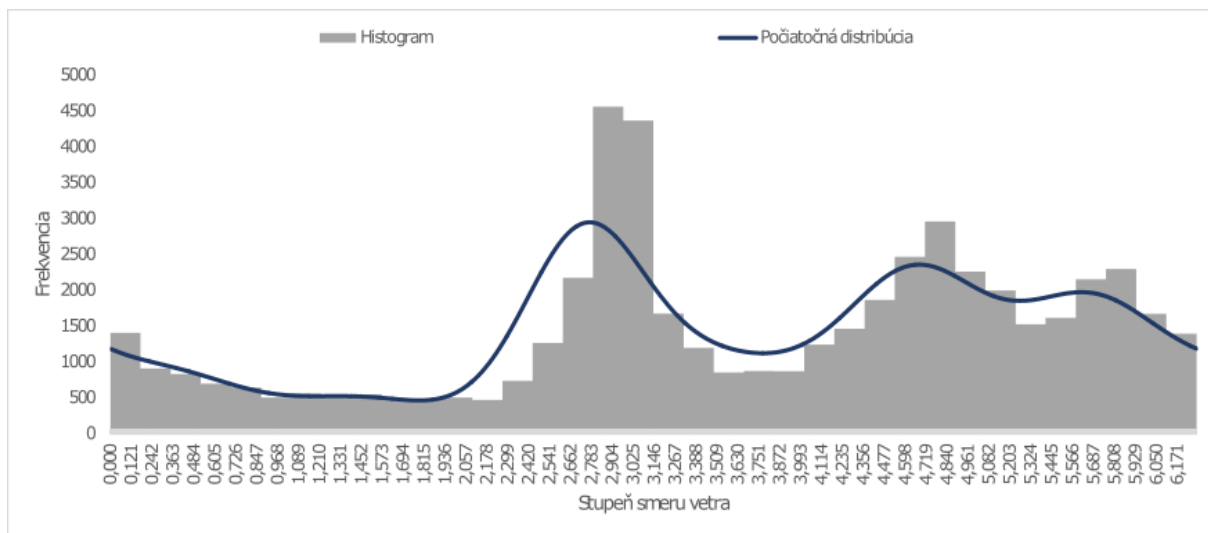
### 6.2.2 Kámen

Dáta pochádzajú z veternej elektrárne Kámen, ktorá je situovaná neďaleko obec Kámen v kraji Vysočina v Českej republike. Fungovať začala v roku 2007 a je tu umiestnená jedna veterná turbína. Všetky parametre počiatočnej zmesi von Misesových distribúcií sú uvedené v tabuľke 3.

Tabuľka 3: Počiatočné parametre jednotlivých von Misesových distribúcií - Kámen

N	$\kappa_j$	$\mu_j$	$\omega_j$
1	5.03865	0.340043	0.09658
2	4.87462	1.46871	0.060005
3	8.16295	2.73716	0.260449
4	4.20677	3.53657	0.129153
5	5.91685	4.65458	0.249622
6	5.776754	5.66954	0.204518

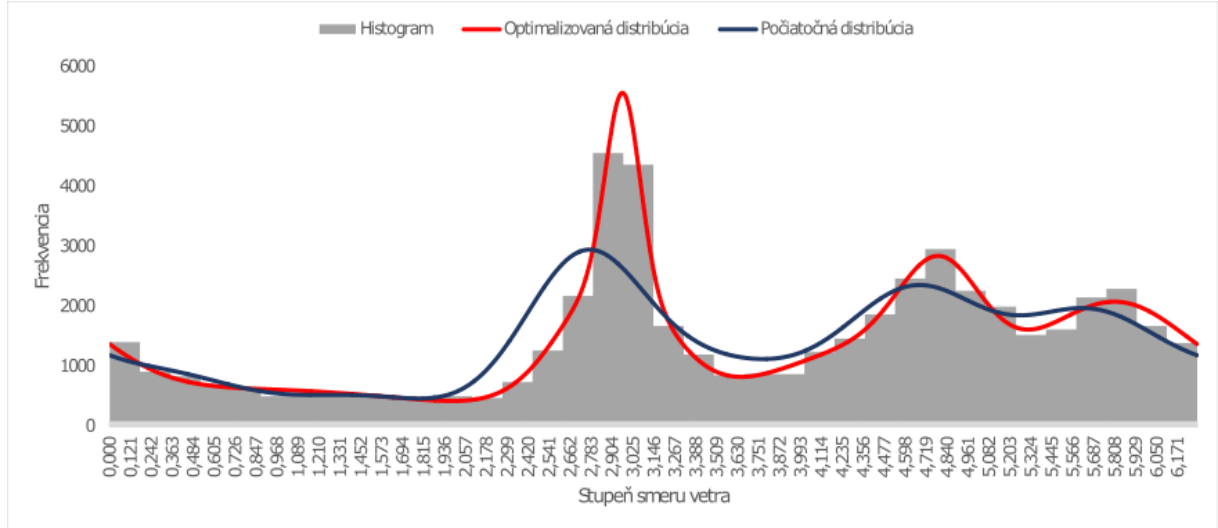
Na obrázku 9 môžeme vidieť normalizovaný frekvenčný histogram spoločne s počiatočnou distribúciou smeru vetra. Rovnako aj na týchto dátach bolo prevedených 30 experimentov a algoritmus bol nastavený na rovnaké parametre ako predtým. Rovnako bol použitý aj test RMSE. Aj pri tomto teste došlo k terminácii algoritmu po dosiahnutí maximálneho počtu iterácií. V tabuľke 4 môžeme vidieť porovnanie výsledkov algoritmu CMA-ES a algoritmu DE. Na obrázku 10 môžeme vidieť frekvenčný histogram spoločne s analyticky odhadnutou distribúciou, a tiež s distribúciou optimalizovanou pomocou algoritmu CMA-ES. Ide o najlepší prípad, kedy sa hodnota  $RMSE = 100,86$ .



Obr. 9: Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Kámen

Tabuľka 4: Výsledky optimalizácie distribúcie smeru vetra pre Kámen

Algoritmus	najlepší	priemer	najhorší
(8,17)-CMA-ES	100,86	139,53	213,53
DE/1/rand	64,77	86,85	136,27



Obr. 10: Normalizovaný frekvenčný histogram s počiatkovou a optimalizovanou distribúciou - Kámen

### 6.2.3 Lipná

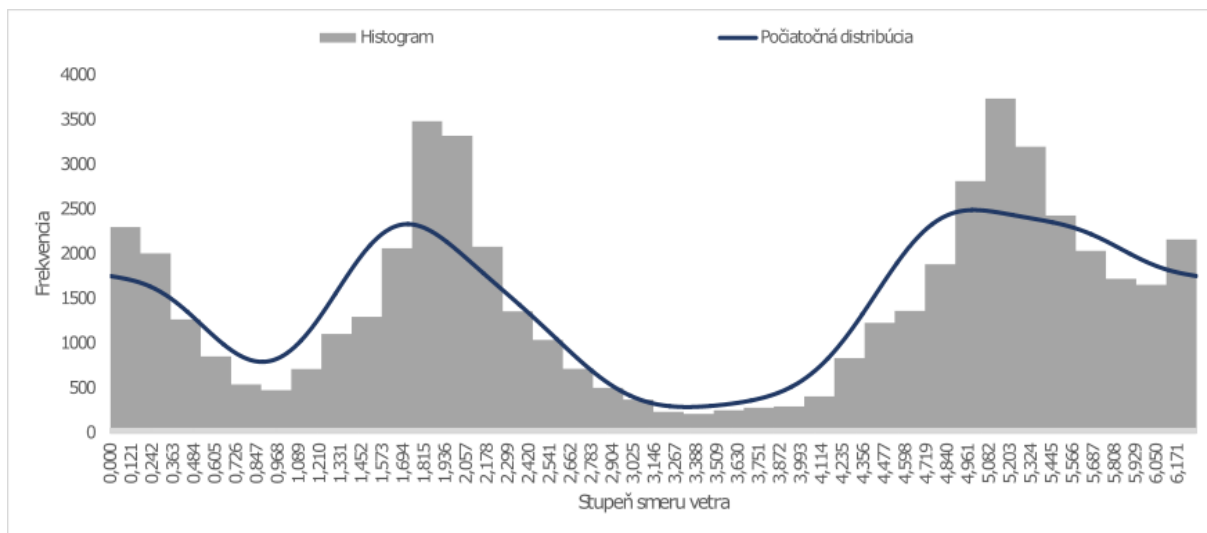
Dáta pochádzajú z elektrárne Lipná, ktorá sa nachádza neďaleko mesta Potštát v okrese Přerov v Olomouckom kraji, ČR. Elektráreň je daná do prevádzky od roku 2008 a nachádza sa v nej jedna veterná turbína. Parametre počiatkovej zmesi von Misesových distribúcií sú uvedené v tabuľke 5.

Tabuľka 5: Počiatkové parametre jednotlivých von Misesových distribúcií - Lipná

N	$\kappa_j$	$\mu_j$	$\omega_j$
1	6.32971	0.271464	0.142647
2	6.4068	1.63788	0.229147
3	6.055	2.36456	0.116318
4	4.76185	3.63703	0.0327544
5	5.6956	4.77268	0.226821
6	4.47652	5.61941	0.252312

Na obrázku 11 je frekvenčný histogram spoločne s počiatkovou distribúciou smeru vetra. Aj na týchto dátach bolo prevedených 30 meraní a pri každom meraní došlo k terminácii po dosiahnutí maximálneho počtu iterácií. Bol použitý test RMSE. V tabuľke 6 môžeme vidieť porovnanie dosiahnutých výsledkov s výsledkami algoritmu DE. Na obrázku 12 je potom najlepší

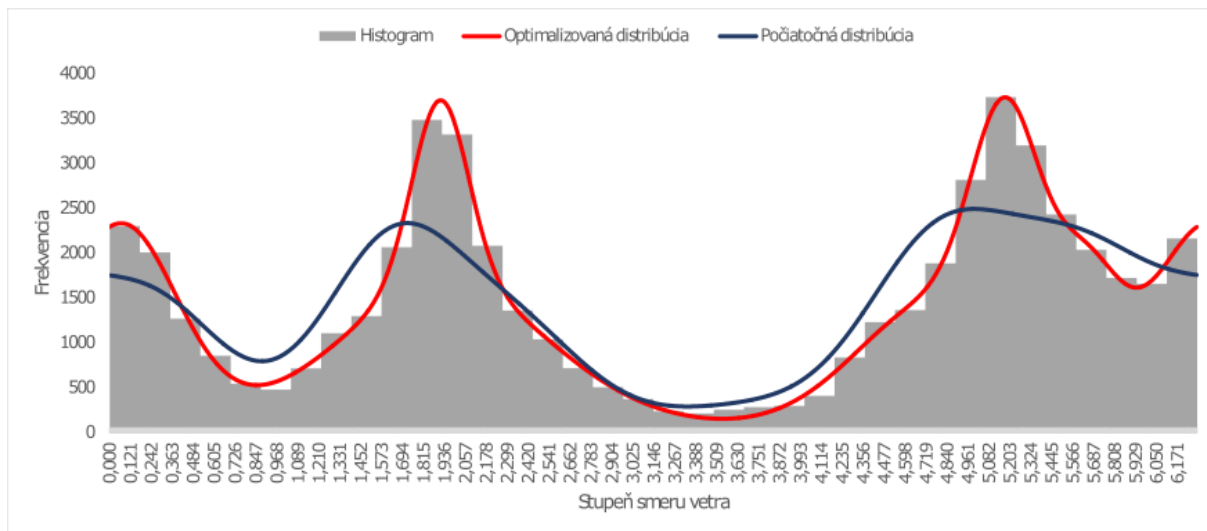
prípád, ktorý bol pri experimentoch dosiahnutý. Je zobrazený frekvenčný histogram spoločne s počiatočnou a optimalizovanou distribúciou. V tomto prípade dosiaholo RMSE hodnotu 51,42.



Obr. 11: Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Lipná

Tabuľka 6: Výsledky optimalizácie distribúcie smeru vetra pre Lipná

Algoritmus	najlepší	priemer	najhorší
(8,17)-CMA-ES	51,42	98,89	169,12
DE/1/rand	49,16	61,33	89,27



Obr. 12: Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Lipná

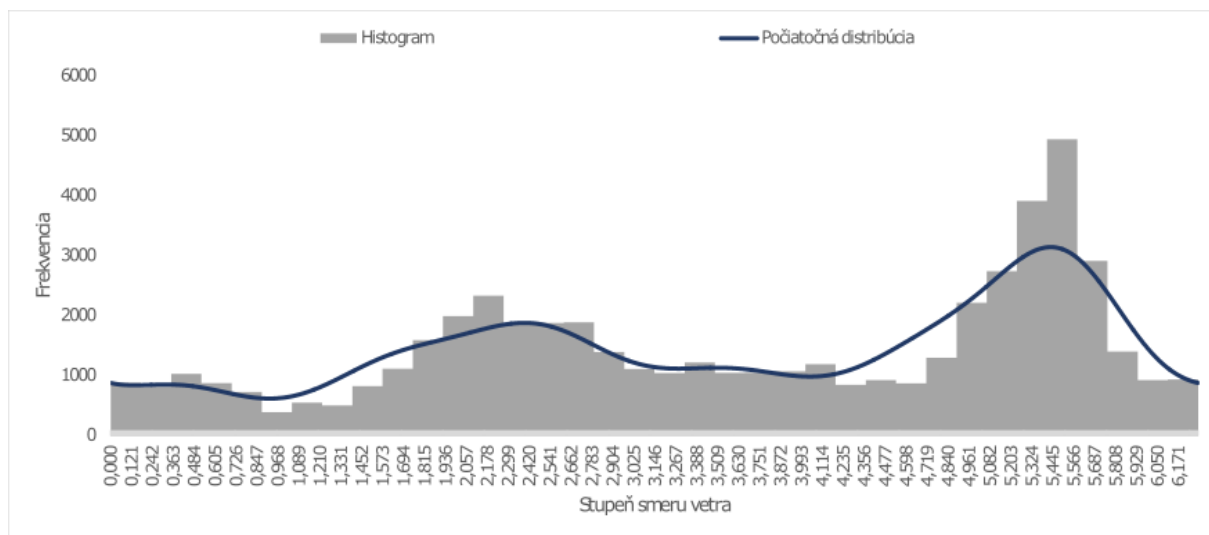
#### 6.2.4 Maletín

Dáta pochádzajú z veternej elektrárne Maletín, ktorá je situovaná v okrese Šumperk v Olomouckom kraji v Českej republike. Bola daná do prevádzky v roku 2008 a je v nej inštalovaná jedna veterná turbína. Všetky parametre počiatočnej zmesi von Misesových distribúcií sú popísané v tabuľke 7.

Tabuľka 7: Počiatočné parametre jednotlivých von Misesových distribúcií - Maletín

N	$\kappa_j$	$\mu_j$	$\omega_j$
1	5.97278	0.383415	0.0902935
2	5.74117	1.62904	0.125751
3	5.32506	2.46736	0.200118
4	4.86776	3.58183	0.126533
5	5.22132	4.76283	0.170078
6	7.06367	5.50493	0.287227

Obrázok 13 zobrazuje normalizovaný dátový frekvenčný histogram spoločne s počiatočným modelom. Na týchto dátach bola prevedená séria 30 experimentov. Fitness hodnota bola vyhodnocovaná pomocou RMSE testu. Pri každom experimente sa algoritmus zastavil po dosiahnutí maximálneho počtu iterácií. V tabuľke 8 je zobrazený najhorší, najlepší a priemerný výsledok dosiahnutý počas experimentov, a tiež porovnanie s algoritmom DE. Na obrázku 14 je zobrazená optimalizovaná distribúcia pomocou algoritmu CMA-ES. V tomto prípade je zobrazená najlepšia hodnota RMSE testu, ktorá mala hodnotu 84,43.

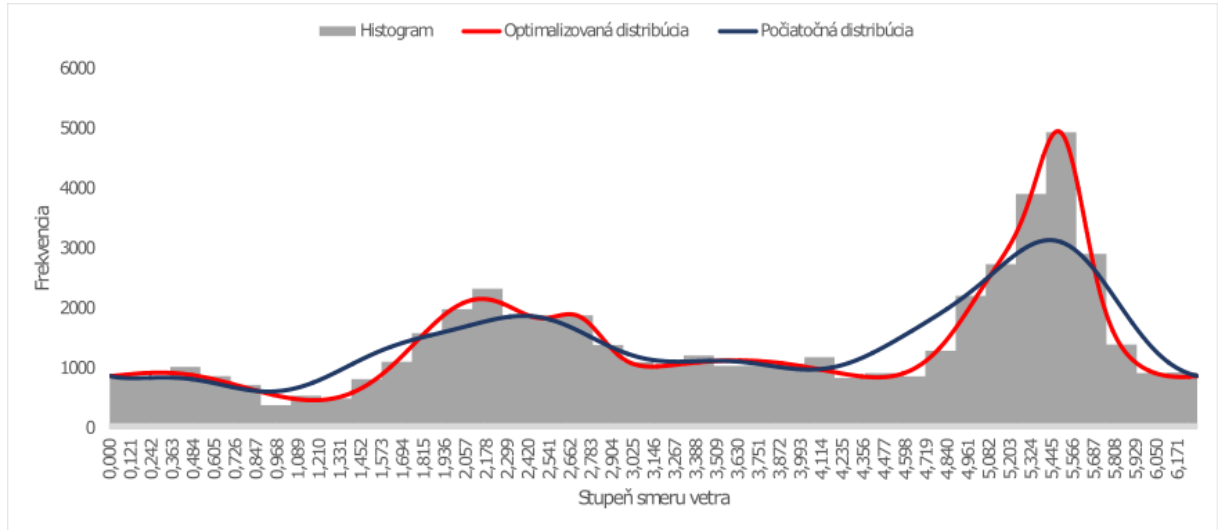


Obr. 13: Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Maletín



Tabuľka 8: Výsledky optimalizácie distribúcie smeru vetra pre Maletín

Algoritmus	najlepší	priemer	najhorší
(8,17)-CMA-ES	84,43	112,12	178,00
DE/1/rand	83,07	96,37	107,10



Obr. 14: Normalizovaný frekvenčný histogram s počiatkovou a optimalizovanou distribúciou - Maletín

### 6.2.5 Veselí u Oder 1

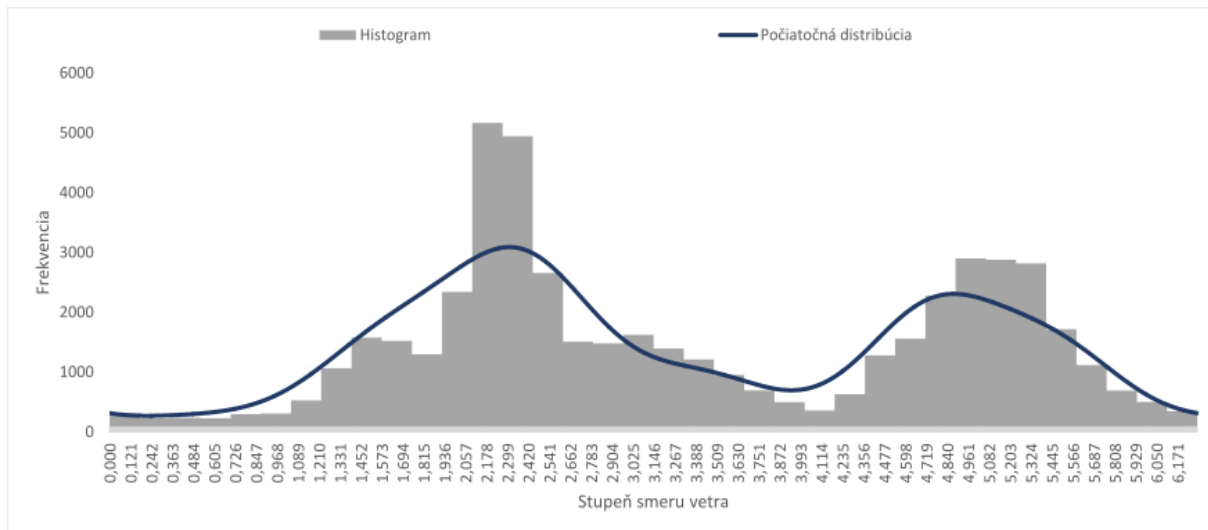
Dáta pochádzajú z veternej elektrárne Veselí u Oder, ktorá je situovaná v obci Odry v okrese Nový Jičín v Moravskoslezskom kraji v Českej republike. Do prevádzky bola daná v roku 2006 a sú v nej umiestnené dve veterné turbíny, pričom tento dataset obsahuje údaje z turbíny 26367. Parametre počiatkových von Misesových distribúcií sú uvedené v tabuľke 9.

Tabuľka 9: Počiatočné parametre jednotlivých von Misesových distribúcií - Veselí 26367

N	$\kappa_j$	$\mu_j$	$\omega_j$
1	4.49868	0.448273	0.0313197
2	5.70516	1.58638	0.162917
3	5.5392	2.38085	0.339841
4	5.99595	3.44649	0.0998865
5	6.41616	4.7541	0.225341
6	6.63737	5.47023	0.1406947

Na obrázku 15 môžeme vidieť normalizovaný frekvenčný histogram a analyticky odhadnutú distribúciu smeru vetra. Aj v tomto prípade bolo prevedených 30 experimentov kvôli stochastickej povahe algoritmu. Parametre algoritmu boli rovnaké ako vo všetkých prípadoch. K terminácii došlo vo všetkých meraniach po dosiahnutí maximálneho počtu iterácií. V tabuľke 10 môžeme

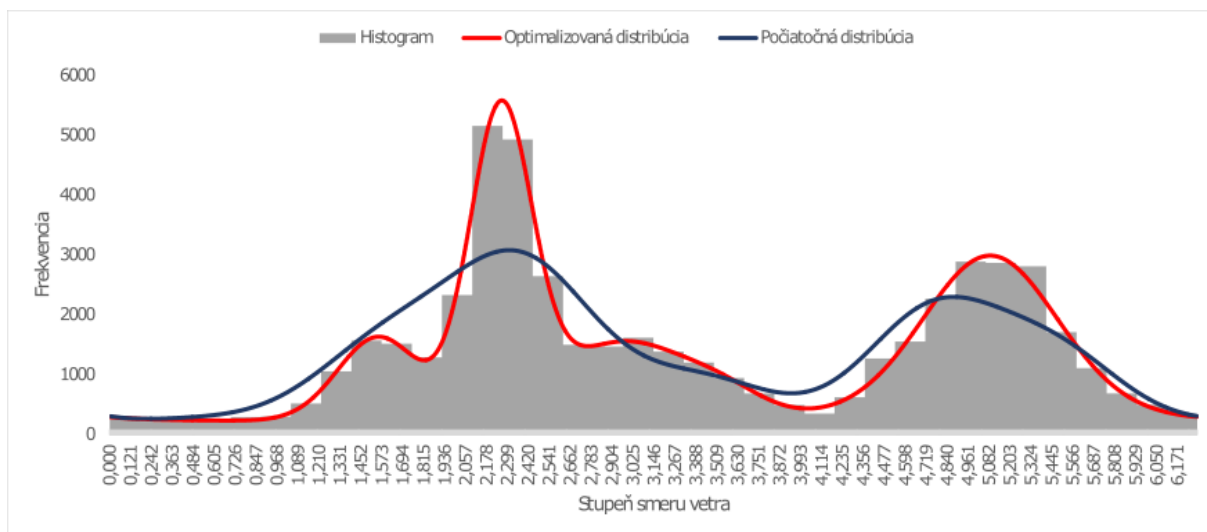
vidieť porovnanie CMA-ES algoritmu a DE algoritmu pri použití testu RMSE. Na obrázku 16 môžeme vidieť normalizovaný frekvenčný histogram s optimalizovanou distribúciou pomocou CMA-ES. Je tu zobrazený najlepší prípad kedy hodnota RMSE bola 81,72.



Obr. 15: Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Veselí 26367

Tabuľka 10: Výsledky optimalizácie distribúcie smeru vetra pre Veselí 26367

Algoritmus	najlepší	priemer	najhorší
(8,17)-CMA-ES	81,72	102,68	169,63
DE/1/rand	62,89	69,06	83,67



Obr. 16: Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Veselí 26367

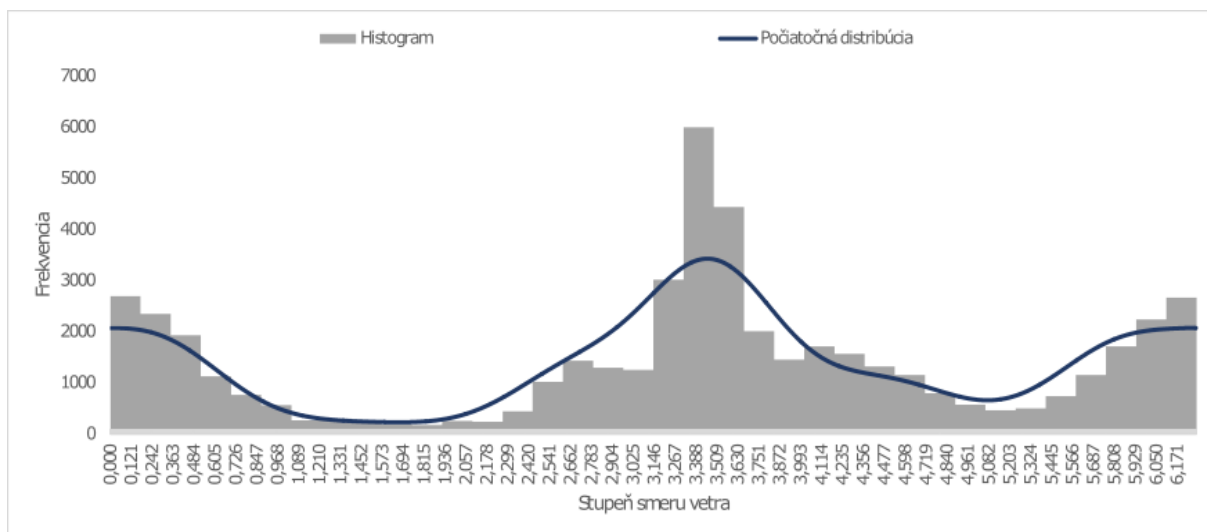
### 6.2.6 Veselí u Oder 2

Dáta pochádzajú z rovnakej elektrárne ako pri predošlom teste, akurát obsahujú merania pre druhú veternú turbínu, ktorá je označená ako 26368. Parametre počiatonej zmesi von Misesových distribúcií sú uvedené v tabuľke 11.

Tabuľka 11: Počiatočné parametre jednotlivých von Misesových distribúcií - Veselí 26368

N	$\kappa_j$	$\mu_j$	$\omega_j$
1	6.47484	0.285048	0.188402
2	4.63227	1.4562	0.0287265
3	7.50908	2.65467	0.113852
4	6.60995	3.47029	0.371297
5	5.82462	4.50327	0.117649
6	6.47734	5.8273	0.180072

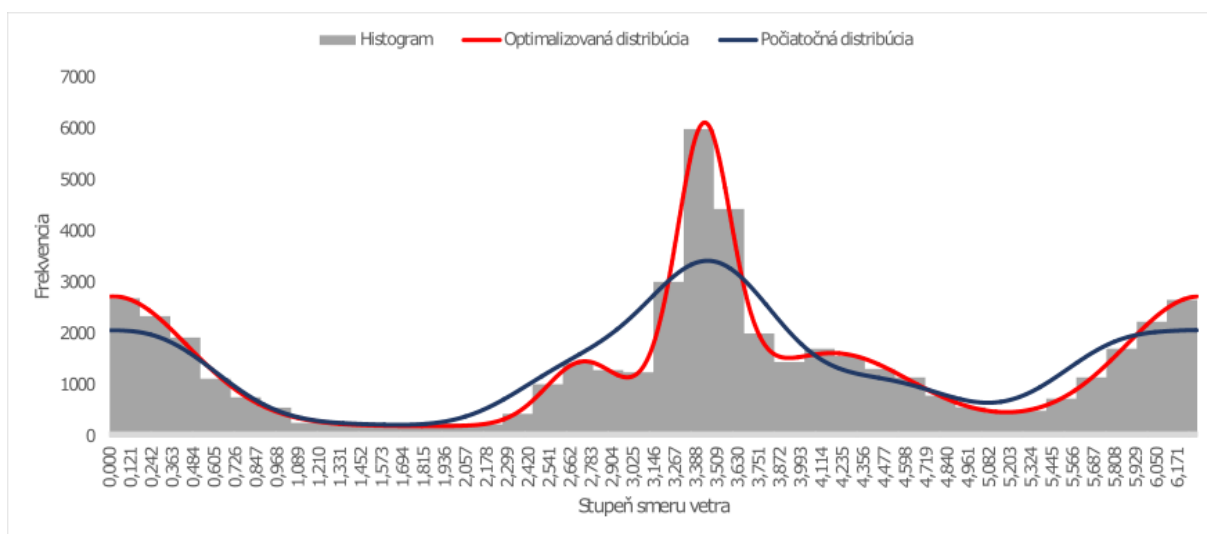
Na obrázku 17 je znázornený normalizovaný frekvenčný histogram spoločne s analytickou distribúciou. Rovnako aj na týchto dátach bolo prevedených 30 meraní, pri každom meraní bola použitá fitness funkcia RMSE. V každom meraní došlo k terminácii po dosiahnutí maximálneho počtu iterácií. V tabuľke 12 môžeme vidieť porovnanie algoritmov CMA-ES a DE. Na obrázku 18 je frekvenčný histogram pre najlepší prípad, kedy sa  $RMSE=44,70$ . Na obrázku je rovnako zobrazená aj optimalizovaná distribúcia pomocou CMA-ES.



Obr. 17: Normalizovaný frekvenčný histogram s počiatočnou distribúciou - Veselí 26368

Tabuľka 12: Výsledky optimalizácie distribúcie smeru vetra pre Veselí 26368

Algoritmus	najlepší	priemer	najhorší
(8,17)-CMA-ES	44,70	61,72	83,68
DE/1/rand	41,82	43,63	48,59



Obr. 18: Normalizovaný frekvenčný histogram s počiatočnou a optimalizovanou distribúciou - Veselí 26368

Výsledky testovania nám ukazujú, že pri použití RMSE testu ako fitness funkcie algoritmus (8,17)-CMA-ES dosahuje v najlepšom prípade rovnaké prípadne porovnateľné výsledky ako algoritmus DE/1/rand. Pri porovnaní priemerných výsledkov však CMA-ES nedosahuje tak dobré

výsledky ako DE. To môže byť spôsobené nastavením oboch algortimov, a tiež rozložením dát v jednotlivých datasetoch. Aj napriek tomu nám však výsledky ukazujú, že navrhnutý CMA-ES algoritmus, môže byť použitý pre optimalizáciu distribúcie smeru vetra, a že distribúcia s optimalizovanými parametrami popisuje reálne dáta lepšie ako analytický odhad. V tabuľke 13 môžeme vidieť porovnanie hodnôt RMSE pre optimalizovaný model a hodnôt RMSE pre analytický model. Z tabuľky môžeme vidieť, že v každom prípade má distribúcia optimalizovaná pomocou algoritmu CMA-ES lepšiu hodnotu RMSE testu ako analytický model.

Tabuľka 13: Porovnanie najlepšieho, priemerného a najhoršieho RMSE optimalizovanej distribúcie s RMSE analytického modelu pre každú lokáciu

Miesto	najlepší CMA-ES (RMSE)	priemer CMA-ES (RMSE)	najhorší CMA-ES (RMSE)	analytický model (RMSE)
Drahany	87,51	133,33	189,14	258,20
Kámen	100,86	139,53	213,53	565,70
Lipná	51,42	98,89	169,12	498,71
Maletín	84,43	112,12	178,00	452,94
Veselí 26367	81,72	102,68	169,63	618,46
Veselí 26368	44,70	61,72	83,68	609,10

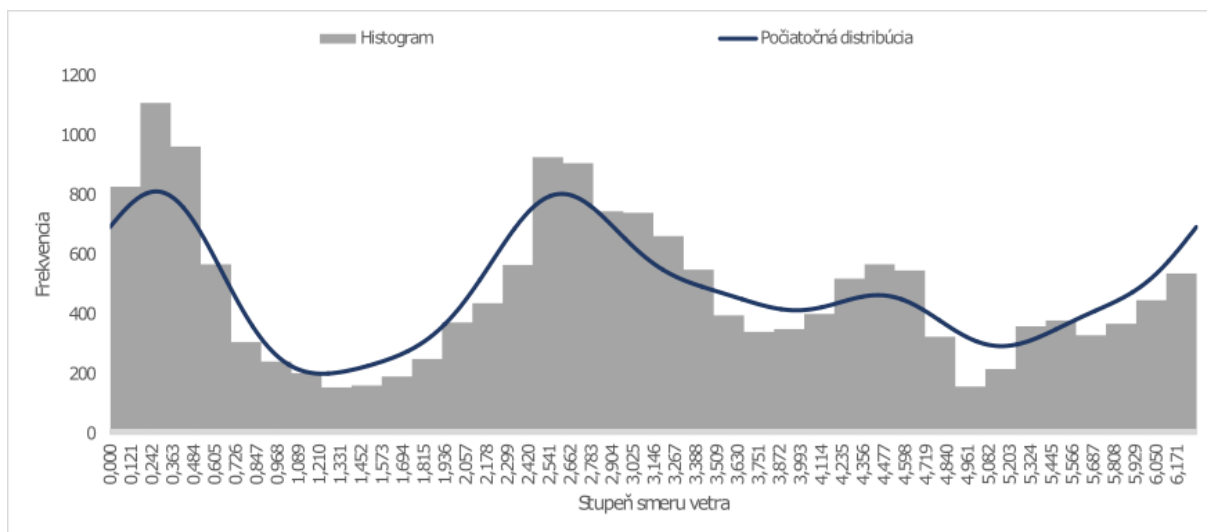
### 6.2.7 Mildred Lake

Dáta pochádzajú z oblasti Mildred Lake, Alberta v Kanade a obsahujú hodinové rozloženie smeru vetra od 1. 1. 2012 do 31. 12. 2013. Všetky parametre počiatočnej zmesi von Misesových distribúcií sú uvedené v tabuľke 14.

Tabuľka 14: Počiatočné parametre jednotlivých von Misesových distribúcií - Mildred Lake

N	$\kappa_j$	$\mu_j$	$\omega_j$
1	7.13198	0.310609	0.230405
2	4.42861	1.56106	0.0783084
3	5.95487	2.57212	0.251946
4	4.57975	3.51352	0.157837
5	6.15393	4.51289	0.136343
6	4.65121	5.71358	0.141513

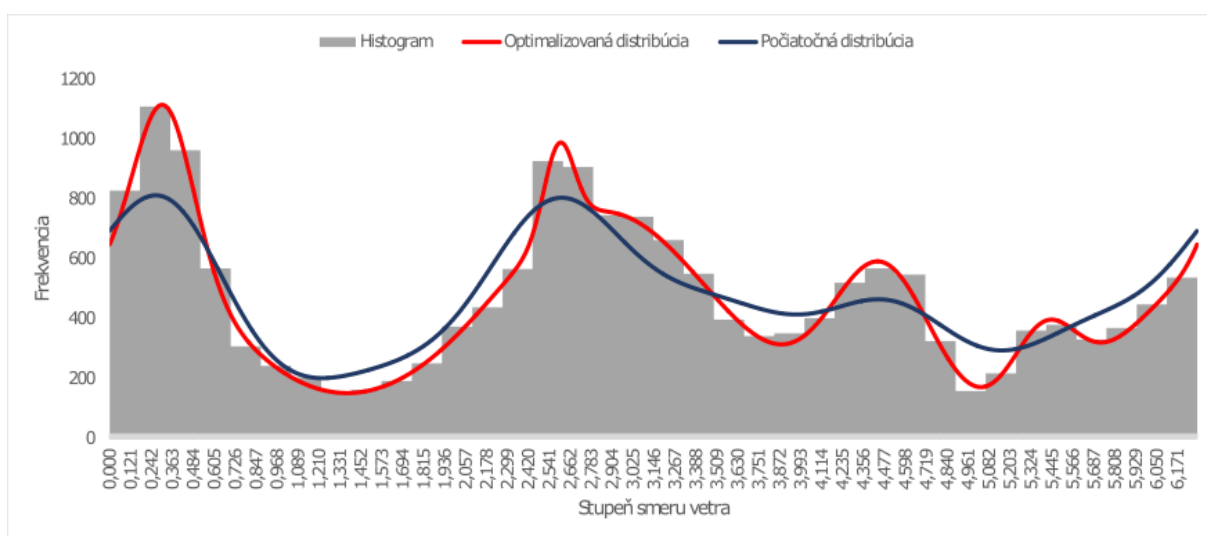
Na obrázku 19 je zobrazený normalizovaný dátový frekvenčný histogram spoločne s počiatočnou distribúciou. Na týchto dátach bolo prevedených 30 experimentov kvôli stochastickému správaniu sa algoritmu CMA-ES. Ako fitness funkcia bol použitý  $\chi^2$  test a to kvôli tomu, aby výsledky mohli byť porovnané s algoritmom PSO, ktorý bol použitý v [15]. Pri každom experimente sa algoritmus CMA-ES zastavil pri dosiahnutí maximálneho počtu iterácií. V tabuľke 15 je zobrazený najlepší, priemerný a najhorší výsledok CMA-ES a jeho porovnanie s PSO. Na obrázku 20 je zobrazený normalizovaný dátový frekvenčný histogram spoločne s počiatočnou a optimalizovanou distribúciou. Obrázok znázorňuje prípad kedy  $\chi^2 = 35,98$ .



Obr. 19: Normalizovaný frekvenčný histogram s počiatkovou distribúciou - Mildred Lake

Tabuľka 15: Výsledky optimalizácie distribúcie smeru vetra pre Mildred Lake

Algoritmus	najlepší	priemer	najhorší
(8,17)-CMA-ES	35,98	151,94	360,25
PSO	112,69	141,64	194,90



Obr. 20: Normalizovaný frekvenčný histogram s počiatkovou a optimalizovanou distribúciou - Mildred Lake

Výsledky experimentov ukázali, že pri použití  $\chi^2$  testu ako fitness funkcie algoritmus (8,17)-CMA-ES dosahuje v najlepšom prípade oveľa lepšie výsledky ako algoritmus PSO. Pri porovnávaní priemerných výsledkov však obidva algoritmy dosahujú porovnateľné výsledky. Výsledky

nám ukázali, že navrhnutý algoritmus CMA-ES môže byť použitý pre optimalizáciu distribúcie smeru vetra, a že dosahuje porovnateľných výsledkov s inými biologicky inšpirovanými algoritmi. V tabuľke 16 môžeme vidieť porovnanie hodnôt  $\chi^2$  pre optimalizovaný model a hodnôt  $\chi^2$  pre analytický model. Môžeme vidieť, že distribúcia optimalizovaná pomocou CMA-ES dosahuje lepšej hodnoty ako analytický model.

Tabuľka 16: Porovnanie najlepšieho, priemerného a najhoršieho  $\chi^2$  optimalizovanej distribúcie s  $\chi^2$  analytického modelu

Miesto	najlepší CMA-ES ( $\chi^2$ )	priemer CMA-ES ( $\chi^2$ )	najhorší CMA-ES ( $\chi^2$ )	analytický model ( $\chi^2$ )
Mildred Lake	35,98	151,94	360,25	620,94

## 7 Záver

Táto práca predstavila novú biologicky inšpirovanú meta-heuristickú metódu pre modelovanie kruhových dát založenú na relatívne novom a populárnom algoritme Covariance Matrix Adaptation - Evolution Strategy. Výpočet ročného rozdelenia smeru vetra je založený na štatistickom rozdelení smerových dát s viacerými prevládajúcimi smermi. Parametre takejto distribúcie nie je možné odhadnúť analyticky. CMA-ES je vhodný pre nájdenie funkcie hustoty pravdepodobnosti zmesi von Misesových distribúcií vďaka svojej schopnosti riešiť problémy s plávajúcou desatinnou čiarkou, a tiež vďaka dobrým výsledkom v oblasti optimalizácie reálnych problémov. Funkcia hustoty pravdepodobnosti zmesi von Misesových distribúcií bola reprezentovaná ako vektor  $\mathbf{v}$ , aby sa dala použiť s algoritmom CMA-ES. Navrhnutý algoritmus bol otestovaný na smerových dátach, ktoré boli zozbierané na 7 rôznych lokalitách. 6 týchto lokalít sa nachádza v Českej republike a jedna v Kanade. Experimenty ukázali, že CMA-ES je schopný vylepšiť počiatočný vektor  $\mathbf{v}$ , ktorý bol analyticky odhadnutý a je možné ho použiť v oblasti modelovania smeru vetra. Model optimalizovaný pomocou CMA-ES dosahuje v najlepšom prípade výsledky porovnateľné s algoritmom DE, avšak v priemere vykazuje o niečo horšie výsledky ako algoritmus DE. V porovnaní s algoritmom PSO, však dosahuje lepších výsledkov. V prípade 6 lokalít z Českej republiky sa na vyhodnotenie či optimalizovaný model pokrýva 100% reálnych dát použil test RMSE. Avšak ani v jednom prípade neboli podmienky testu splnené. V prípade lokality z Kanady bol použitý  $\chi^2$  test a hoci optimalizovaný model je lepší ako predchádzajúci stále nespĺnil podmienky testu. Preto zo štatistického hľadiska optimalizované modely nepokrývajú dáta. Avšak po vizuálnom posúdení výsledných histogramov, môžeme vidieť, že došlo k výraznému zlepšeniu oproti analytickým modelom.

V tejto oblasti je veľa príležitostí pre budúce práce. Použitie iných fitness funkcií (napr.  $R^2$ ) rovnako ako použitie iných optimalizačných metód (napr. EDA) pre modelovanie smerových dát vetra. Ďalším smerom kam by sa mohol výskum uberať je algoritmus, ktorý by automaticky vybral počet distribúcií, ktoré sa musia pomiešať, aby reprezentovali dáta. Z pohľadu použitia môže byť aplikácia použitá na modelovanie iných rozličných typov kruhových dát.

Bc. Michal Šamaj



## Literatúra

- [1] Mooney J. A. a Jolliffe P. J., *Fitting mixtures of von Mises distributions: a case study involving sudden infant death syndrome*, Computational Statistics & Data Analysis 41, strany 505-513, 2003.
- [2] Calderara S., Cucchiara R. a Prati A., *Detection of Abnormal Behaviors Using a Mixture of Von Mises Distributions*, Proceedings of the 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, AVSS 2007, strany 141-146, Washington DC, 2007.
- [3] Vejmelka M., Musilek P. Paluš M. a Pelikán E., *K-means clustering for problems with periodic attributes*, International Journal of Pattern Recognition and Artificial Intelligence 23 (4), strany 721-743, 2009.
- [4] Jung S. a Kwon S. D., *Weighed error functions in artificial neural networks for improved wind energy potential estimation*, Applied Energy 111, strany 778-790, 2013.
- [5] Fried L., *Global Wind Statistics 2014*, Global Wind Energy Council.
- [6] *Wind Energy Basic* online: [http://www.awea.org/faq/wwt\\_basics.html](http://www.awea.org/faq/wwt_basics.html) [marec 2016]
- [7] Carta, J. A., Bueno, C. a Ramírez, P., *Statistical modelling of directional wind speeds using mixtures of von mises distributions: Case study*, Energy Conversion and Management 49 (5), strany 897-907, 2008.
- [8] Ettoumi F., Sauvageot H. a A. Adane A., *Statistical bivariate modelling of wind using first-order Markov chain and Weibull distribution*, Renew. Energy, No. 28, strany 1787-1802, 2003.
- [9] Davis M., *A new thermal rating approach: The real time thermal rating system for strategic overhead conductor transmission lines, part I: General description and justification of the real time thermal rating system*, IEEE Trans. On Power Apparatus and Systems, No.96, strany 803-809, 1977.
- [10] Allen, C. T., Young G. S. a Haupt S. E., *Improving pollutant source characteriyation by better estimating wind direction with genetic algorithm*, Atmospheric Enviroment 41 (11), strany 2283-2289, 2007.
- [11] Koch W., *Directional analysis of SAR images aiming at wind direction*, IEEE Transactions on Geosience and Remote Sensing 42 (4), strany 702-710, 2004.
- [12] Erdem E. a Shi J., *ARMA based approaches for forecasting the tuple of wind speed and direction*, Applied Energy 88 (4), strany 1405-1414, 2011.

- [13] de la Rosa J. J. G., Pérez A. A., Salas J. C. P., Leo J. G. R. a Munoz A. M., *A novel inference method for local wind condition using genetic fuzzy systems*, Renewable Energy 36 (6), strany 1747-1753, 2011.
- [14] Heckenbergerova J., Musílek P., Mejznar J. a Vancura M., *Estimation of wind direction distribution with genetic algorithm*, CCECC, strany 1-4, 2013.
- [15] Heckenbergerova J., Musílek P. a Kromer P., *Optimization of wind direction distribution parameters using particle swarm optimization*, Afro-European Conference for Industrial Advancement, strany 15-26, 2014.
- [16] Shamshirband S., Iqbal J., Petkovic D. a Mirhashemi M. A., *Survey of four models of probability density functions of wind speed and directions by adaptive neuro-fuzzy methodology*, Advances in Engineering Software 76, strany 148-153, 2014.
- [17] Risken H., *The Fokker–Planck Equation*, Springer, ISBN 978-3-540-61530-9, 1989.
- [18] Mardia Kantilal; Jupp, Peter E., *Directional Statistics*, Wiley, ISBN 978-0-471-95333-3, 1999.
- [19] Fisher, Nicholas I., *Statistical Analysis of Circular Data*, Cambridge, New York, 1993.
- [20] Abramowitz M. a Stegun I. A., *Handbook of Mathematical Functions*, National Bureau of Standards, 1964. ISBN 0-486-61272-4.
- [21] Jammalamadaka S. Rao a Sengupta A., *Topics in Circular Statistics*, World Scientific Publishing Company, ISBN 978-981-02-3778-3, 2001.
- [22] L. J. Fogel, A. J. Owens, a M. J. Walsh, *Artificial Intelligence through Simulated Evolution*, New York: John Wiley, 1966.
- [23] J. H. Holland, *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, 1975.
- [24] I. Rechenberg, *Evolutionstrategie: Optimierung Technischer Systeme nach Prinzipien des Biologischen Evolution*, Fromman-Holzboog Verlag, Stuttgart, 1973.
- [25] H.-P. Schwefel, *Numerical Optimization of Computer Models*, John Wiley and Sons, New-York, 1981.
- [26] Z. Michalewicz, *Genetic Algorithms + Data Structures - Evolution Programs*, Springer, 1996. ISBN 3540606769.
- [27] Barricelli Nils Aall, *Numerical testing of evolution theories : Part I Theoretical introduction and basic tests*, Acta Biotheoretica 16 (1-2), strany 69-98, 1962.

- [28] Fraser A. S., *Monte Carlo analyses of genetic models*, Nature 181 (4603), strany 208-209, 1958.
- [29] Koza, John R., *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, ISBN 0-262-11170-5, 1992.
- [30] Jamshidi M., *Tools for intelligent control: fuzzy controllers, neural networks and genetic algorithms*, Philosophical Transactions of the Royal Society A 361 (1809), strany 1781-1808, 2003.
- [31] Hansen N. a Ostermeier A., *Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation*, Proceedings of the IEEE International Conference on Evolutionary Computation, strany 312–317, 1996.
- [32] Hansen N. a Ostermeier A., *Convergence properties of evolution strategies with the derandomized covariance matrix adaptation: The  $(\mu/\mu_I, \lambda)$ -ES*, EUFIT'97, 5th European Congress on Intelligent Techniques and Soft Computing Proceedings, strany 650-654, 1997.
- [33] Hansen N. a Kern S., *Evaluating the CMA Evolution Strategy on Multimodal Test Functions*, Eighth International Conference on Parallel Problem Solving from Nature PPSN VIII, strany 282-291, Springer, Berlin, 2004.
- [34] Hansen, N., *The CMA evolution strategy: a comparing review, Towards a new evolutionary computation. Advances on estimation of distribution algorithms*, Springer, strany 1769–1776, 2006.
- [35] Auger, A. a N. Hansen, *A Restart CMA Evolution Strategy With Increasing Population Size*, Proceedings of the IEEE International Conference on Evolutionary Computation, 2005.
- [36] Akimoto Y., Y. Nagata, I. Ono a S. Kobayashi, *Bidirectional Relation between CMA Evolution Strategies and Natural Evolution Strategies*, Parallel Problem Solving from Nature, PPSN XI. Springer, strany 154–163, 2010.
- [37] Glasmachers T., T. Schaul, Y. Sun, D. Wierstra a J. Schmidhuber, *Exponential Natural Evolution Strategies*, Genetic and Evolutionary Computation Conference GECCO. Portland, OR, 2010.
- [38] Nelder John A. a Mead R. *A simplex method for function minimization*. Computer Journal 7, strany 308-313, 1965.
- [39] Broyden C. G., *The convergence of a class of double-rank minimization algorithms*, Journal of the Institute of Mathematics and Its Applications 6, strany 76-90, 1970.
- [40] Powell M. J. D., *The NEWUOA software for unconstrained optimization without derivatives*, Department of Applied Mathematics and Theoretical Physics, Cambridge University, 2004.

- [41] Igel C., Sutton N. a Hansen N., *A Computational Efficient Covariance Matrix Update and a (1+1)-CMA for Evolution Strategies*, Proceedings of the Genetic and Evolutionary Computation Conference (GECCO), strany 453-460, ACM Press, 2006.
- [42] Igel C., Hansen N. a Roth S., *Covariance Matrix Adaptation for Multi-objective Optimization*, Evolutionary Computation 15 (1), strany 1-28, MIT Press, 2007.
- [43] Jasterbski G. A. a Arnold D. V., *Improving Evolution Strategies through Active Covariance Matrix Adaptation*, Proceedings of IEEE World Congress on Computational Intelligence, strany 9719-9726, 2006.
- [44] Hansen N., *The CMA Evolution Strategy: A Tutorial*, online dostupné z: <https://www.lri.fr/~hansen/cmatutorial.pdf> [marec 2016]
- [45] Pearson K., *On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling*, Philosophical Magazine Series 5 50 (302): strany 157–175.
- [46] Colin Cameron A., Windmeijer Frank A. G., Gramajo H., Cane D. E. a Khosla C., *An R-squared measure of goodness of fit for some common nonlinear regression models*, Journal of Econometrics 77 (2), strany 1790-1792, 1997.
- [47] *A Dictionary of Physics (6 ed.)*, Oxford University Press, 2009. ISBN 9780199233991.
- [48] Zelinka I., *Biologicky inspirované výpočty*, online dostupné z: <http://dataanalysis.vsb.cz/data/Vyuka/BIA/BIV-AUI.pdf> [marec 2016]
- [49] Obrázok dostupný z: [https://en.wikipedia.org/wiki/Von\\_Mises\\_distribution](https://en.wikipedia.org/wiki/Von_Mises_distribution) [marec 2016]
- [50] Obrázok dostupný z: <https://en.wikipedia.org/wiki/CMA-ES> [marec 2016]
- [51] Obrázok dostupný z: [https://en.wikipedia.org/wiki/Multivariate\\_normal\\_distribution](https://en.wikipedia.org/wiki/Multivariate_normal_distribution) [marec 2016]

## A Obsah priloženého CD

Obsah priloženého CD:

- Testovacia aplikácia `WindTool.exe`
- Zložka so zdrojovými kódmi `/sourceCode/`
- Zložka s testovacími dátami `/testData/`
- Návod k aplikácii `ReadMe.txt`
- Textová časť diplomovej práce vo formáte PDF `DP_SAM0036.pdf`